

This chapter presents the criteria developed by EPA as a means for selecting acceptable detection and quantitation limit approaches for use in Clean Water Act (CWA) programs. These criteria reflect EPA's careful consideration of the issues identified and discussed in Chapter 3. A total of six criteria were established, and are discussed in Sections 4.1 - 4.6. Table 4-1 at the end of this chapter summarizes the relationship between each issue discussed in Chapter 3 and the criteria discussed in Sections 4.1 - 4.6.

4.1 Criterion 1

Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

The concept of scientific validity is widely accepted but loosely defined. For the purposes of this evaluation, a detection/quantitation approach or methodology will be considered scientifically valid if it meets the following conditions:

- It can be (and has been) tested,
- It has been subjected to peer review and publication,
- The error rate associated with the approach or methodology is either known or can be estimated,
- Standards exist and can be maintained to control its operation (i.e., it is supported by well-defined procedures for use), and
- It has attracted (i.e., achieved) widespread acceptance within a relevant scientific community.

While EPA acknowledges that other measures could be established to demonstrate scientific validity, EPA has adopted the conditions cited because they reflect those discussed by the U.S. Supreme Court as considerations pertaining to assessments of scientific validity when considering the admissibility of expert scientific testimony¹. EPA believes that considerations discussed by the Court as necessary to demonstrate the scientific validity of an expert's reasoning or methodology are equally valid for demonstrating the scientific validity of a detection/quantitation approach.

4.2 Criterion 2

Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

As discussed in Chapter 3 of this Assessment Document, the detection and quantitation limit(s) for an analyte in an analytical method can be established from a single-laboratory study, multiple single-laboratory studies, or an interlaboratory study. Historical methods developed by EPA under Clean Water Act programs, and nearly all methods developed by EPA under Safe Drinking Water Act programs, were developed by EPA's research laboratory in Cincinnati, Ohio. In the course of method development, this single laboratory established detection and quantitation limits. In many instances, these detection and quantitation limits were found to be unrealistic, in that they could not be achieved in many non-research laboratories. However, with time laboratory and method performance, as well as analytical instrumentation improved, making detection and quantitation limits more easily achievable in nearly all laboratories. Therefore, the difficulty created was in initial application of the research methods.

¹*Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993) and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999)

In recent years, EPA's Office of Science and Technology has used single-laboratory studies to develop an initial estimate of the detection and quantitation limit for a new or modified method, and has verified these limits in interlaboratory studies or by conducting additional single-laboratory studies in other laboratories.

Voluntary consensus standards bodies (VCSBs) such as ASTM International have historically used interlaboratory studies to establish method performance. Over the past 5 to 10 years, ASTM International has been developing interlaboratory and single-laboratory approaches for detection and quantitation. Whereas the single-laboratory studies at EPA's research laboratory in Cincinnati produce the lowest detection and quantitation limits, approaches such as those published by ASTM International gather all sources of variability to produce the highest detection and quantitation limits. A realistic expectation of method and laboratory performance likely lies somewhere in between.

As noted in Section 3.2.2 of this Assessment Document, laboratory and method performance can be affected by the use of performance criteria that serve as prediction or tolerance limits. Examples of such criteria include measures to demonstrate that a laboratory is producing accurate results at a concentration of interest (i.e., analysis of reference standards or spiked samples), measures to demonstrate that results are not biased by contamination (i.e., analysis of blanks), and measures to demonstrate that the laboratory can achieve the sensitivity required to reliably detect pollutants at low concentrations (i.e., at the detection limit). It is likely that laboratory performance will improve (and variability will be lower) when laboratories are required to meet specified performance criteria in order to report results.

A further consideration concerning routine variability is the means for rejection of outliers. True outliers can occur in laboratory data and some means of resolving outlier issues must be included. Statistical procedures are available for the identification of candidate outlier values. Once a candidate outlier has been identified, evaluation of the value from a chemical analytical perspective (e.g., some procedural error or quality control error has occurred) should be the basis of exclusion of the value from a data set. In cases where no cause for the outlier has been identified, it may be reasonable to reject an outlier on statistical grounds, but every effort should be made to justify the exclusion on technical grounds.

In examining each approach against this criterion, EPA will evaluate whether the approach can be used to provide a realistic expectation of laboratory performance. As part of this assessment, EPA will examine the sources of variability captured by the approach, and the degree to which the statistics that underlie the approach realistically reflect these sources of variability.

4.3 Criterion 3

Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

Any approach or procedure should be simple, complete, and cost-effective to implement (i.e., it should be reliable and "laboratory-friendly"). The laboratories that can be expected to use detection or quantitation procedures will range from large laboratories and laboratory chains with a wide range of technical capabilities, to "mom and pop" laboratories operated by one or a few people with a limited set of statistical skills. If a procedure is complicated, it will be prone to error in its use. Similarly, if a procedure requires investment of extensive resources that cannot be billed to the client, laboratories will have a disincentive to use the procedure. Therefore, if the Agency wishes to encourage the development and use of innovative techniques that improve measurement performance or lower measurement costs, the Agency must consider practicality and affordability as significant, if not equal, considerations to scientific validity.

After evaluating each of the issues discussed in Chapter 3 of this document, EPA concluded that successful implementation of CWA programs depends on the ability of laboratories to easily and affordably:

1. demonstrate that a method works in a particular matrix at the levels of concern,
2. characterize improvements in measurement capabilities in terms of measurement sensitivity, and
3. characterize the sensitivity of new methods.

A matrix effect is an interference in a measurement that is caused by substances or materials in the sample other than the analyte of interest that are not removed using the procedures in the method or other commonly applied procedures. In the context of detection and quantitation, matrix effects may manifest themselves by precluding measurements at levels as low as could be measured were the interference not present. From a practical perspective, it is not possible to test the sensitivity of each new method in every possible matrix in which it may be used. At a minimum, it is unlikely that EPA or any other organization could possibly identify and obtain samples of every matrix to which the method might be applied, and even if such a feat were possible, the cost and logistics of doing so would be prohibitive. The situation for characterizing matrix effects on analytical sensitivity is similar to the situation for characterizing matrix effects on measurement performance at higher concentration levels. In the latter case, EPA typically uses one or more spiked reference matrices (e.g., reagent water, sand, diatomaceous earth) to establish QC acceptance criteria for real-world matrix samples that are spiked with the analyte of interest at a mid-to-high concentration. Each analytical method includes QC acceptance criteria for such matrix spikes, along with a suite of quality control requirements designed to verify that failures are attributable to the matrix rather than to an analytical system that is out of control. EPA prefers to identify a similar concept that allows for characterization of measurement sensitivity in representative matrices and that is supported by a simple, cost-effective procedure that would allow individual laboratories to evaluate, on an as-needed basis, the effects of specific matrices on measurement sensitivity. Because methods approved at 40 CFR part 136 already contain a suite of quality control procedures and QC acceptance criteria, EPA believes that it is not necessary to verify method sensitivity in each and every batch of each and every matrix analyzed. Rather, such testing could be done only on an as-needed basis when it is suspected that matrix interferences may preclude reliable measurements at low levels.

Another consideration is that measurement capabilities generally improve over time. This is attributable to a variety of factors, including:

1. increased staff experience with a given technique,
2. technological upgrades or improvements in the instrumentation used for analysis, and
3. development of new instrumentation or techniques that improves sensitivity, precision, or bias.

In each case, the improvements may not be observed across the entire laboratory community. In the case of increased staff experience, for example, it is obvious that a laboratory that specializes in one type of analysis, such as low-level mercury measurements, will develop greater experience than a laboratory that rarely performs this measurement. Likewise, it is easy to see how one or a few laboratories that concentrate their business on a particular type of analysis might be willing to invest significant resources in new or upgraded equipment to improve performance, whereas laboratories that rarely perform such analyses would not find such upgrades to be cost-effective.

Improvements in measurement capability, including the development of new methods, may create a dynamic decision-making process, in that measurements at lower levels may allow EPA and states to identify previously undetected pollutants. Such situations offer a means for monitoring and controlling (i.e., regulating) the discharge of previously unregulated, but harmful, pollutants. Therefore, it is in the best interest of the environment for EPA to encourage the development and use of improved environmental analysis procedures and equipment.

In evaluating this criterion, EPA will favor affordable and easy-to-use approaches and procedures that allow analysts in a single laboratory to 1) determine matrix-specific variations when necessary, based on realistic data, and 2) demonstrate lower detection and quantitation limits associated with improvements in their measurement capabilities. Procedures for establishing the sensitivity of new methods or improved measurement capabilities must be practical enough to encourage such development. These procedures should specify the spiking level at which measurements are to be made and the corrective action to be taken if the resulting detection or quantitation limit is inconsistent with the data from which it is derived.

4.4 Criterion 4

Criterion 4: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.

Any approach to establishing levels at which detection decisions are made should be capable of providing regulators, the regulated community, and data users with a high level of confidence that a pollutant reported as being present really is present. Historically, nearly every approach to making detection decisions has set the criterion for detection at 99 percent confidence (i.e., the lowest level at which a pollutant will be detected with a probability of 99 percent). This criterion results in the probability of a false positive (i.e., that a pollutant will be stated as being present when it actually is not [a Type I error]) of one percent.

In evaluating this criterion, EPA will favor approaches and procedures that reflect routine analytical conditions in a well-operated laboratory. That is, the procedure must be capable of generating a detection level when the substance of interest is not present in a blank and/or when instrument thresholds are adjusted for routine operation.

4.5 Criterion 5

Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.

Measurement capabilities among laboratories vary depending on a number of factors, including, but not limited to, instrumentation, training, and experience. Similarly, measurement capabilities among different analytical methods vary depending on a number of factors, including the techniques and instrumentation employed and the clarity of the method itself.

Historical approaches to recognizing laboratory capabilities in establishing detection and quantitation limits have varied between two extremes of establishing the limit in a state-of-the-art research laboratory to reflect the lowest possible limit that can be achieved, and establishing the limit based on statistical confidence intervals calculated from a large number of laboratories with varying levels of experience, instrumentation and competence. Generally, use of the former has been employed to serve as a goal or performance standard to be met by other laboratories, whereas use of the latter treats the limit, not as a performance standard that needs to be met by each laboratory, but rather as a characterization of the performance of the capabilities of a population of laboratories at the time of method development.

Historical approaches to recognizing method capabilities also have varied between those that allow the error expressed as relative standard deviation, or RSD among low-level measurements to vary, depending on the capabilities of the method, and those that fix this error (RSD) at a specific level.

Initially, Criterion 5 stated that the *“quantitation limit should identify a concentration at which the reliability of the measured result is consistent with the capabilities of the method when a method is performed by experienced staff in a well-operated laboratory.”* Reviewers from within EPA questioned the criterion’s implication that measurements below a quantitation limit could be considered unreliable. A similar concern was expressed by one of the peer reviewers charged with evaluating EPA’s assessment and an earlier draft of this Assessment Document. This reviewer noted that:

“almost all implementations of limits of quantitation have nothing to do with whether the measurements are actually quantitative,” and that “any level at which the instrument can be read, and at which there is a reliably estimated standard deviation is a level at which quantitation is possible” (Rocke, 2002)

The peer reviewer suggested that Criterion 5 might be rewritten as:

“the quantitation limit should identify a concentration at which the instrument yields a measurable signal at least 99% of the time, and which is no smaller than the detection level. Such a quantitation limit will often be the same as the detection level.”

EPA agrees that this is a valid perspective, in that if the pollutant is identified and the analytical system produces a result, quantitation occurs. Although this interpretation of a quantitation limit has validity, implementation of such an approach would require that all values generated by an analytical system be reported, along with an estimate of the uncertainty associated with each value (e.g., the “reliably estimated standard deviation” mentioned by the peer reviewer). As noted in Section 2.3.4, several organizations, including the European Union, are developing procedures for estimating the uncertainty associated with measured results. If successful, such an approach would eliminate many of the data censoring concerns discussed in Section 3.3.2. Given the difficulty in achieving consensus on an appropriate means of establishing a detection limit, however, EPA believes that it would also be difficult, to obtain consensus on an appropriate means for estimating the uncertainty associated with each result measured on each environmental sample. In addition, analytical chemists have used and believe that they understand a quantitation limit to mean the lowest concentration at which an analyte can be identified and determined with some degree of certainty.

Therefore, EPA prefers to monitor developments by the EU and others on this subject, and if appropriate, re-evaluate this issue if and when it becomes widely accepted by the laboratory, regulatory, and regulated communities. In the meantime, EPA believes that the traditional approach of defining a quantitation limit at some level above the detection limit provides a data user with a reasonable degree of confidence in the measured value without requiring that individual estimates of uncertainty be developed and reported. Criterion 5 reflects this belief.

EPA will evaluate various approaches against this criterion by examining the ease of adjustment of the RSD or other performance measures in the context of the measurement capability of the laboratory or the need to adjust the measurement error to allow for environmental decisions. In evaluating the approaches, EPA will give preference to those approaches that strike a reasonable balance between using either state-of-the art laboratories or a highly varied community of laboratories to establish quantitation limits.

4.6 Criterion 6

Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.

The Clean Water Act requires EPA to conduct, implement, and oversee a variety of data gathering programs. As noted in Section 3.2 of this Assessment Document, these programs include, but are not limited to:

- Survey programs to establish baselines and monitor changes in ambient water quality,
- Screening studies to identify emerging concerns and establish the need for more in-depth assessment,
- Effluent guideline studies to establish technology-based standards for the control of pollutants in wastewater discharges,
- Toxicity and environmental assessment studies to establish water quality-based standards for the control of pollutants in wastewater, and
- Risk assessment studies designed to characterize and evaluate human health and environmental risks associated with various water body uses.

In addition, EPA needs to apply a detection limit or quantitation limit approach to permitting, compliance monitoring, and other uses of the 40 CFR part 136 methods. These applications include:

- Permitting,
- Ambient and effluent compliance monitoring under NPDES and the pretreatment program,
- Ambient and effluent compliance monitoring under state and local programs,
- Quality control in analytical laboratories, and
- Method promulgation.

In theory, EPA could evaluate each of these applications independently and identify a detection and quantitation limit approach that is best suited to each application. However, doing so could potentially result in the need for up to 10 different detection and/or quantitation limit approaches. EPA believes that this would increase confusion, increase record keeping burdens, and increase laboratory testing burdens. For these reasons, EPA believes it is desirable to adopt a single pair of related detection and quantitation procedures that can be used to address all Clean Water Act applications.

EPA also believes that 1) it is unrealistic to expect other organizations, such as the U.S. Geological Survey, the Food and Drug Administration, ASTM International, AOAC-International, etc., to adopt and standardize on the approach selected by EPA for its use in CWA programs, and 2) it is desirable to allow use of approaches and methods developed by these and other organizations to be used in CWA programs. The inclusion of such approaches and methods provides the stakeholder community with increased measurement options that may help reduce measurement costs or improve measurement performance for specific situations. This approach is consistent with EPA's movement towards a performance-based measurement system (PBMS) and with the intent of the National Technology Transfer and Advancement Act (NTTAA). Therefore, although EPA prefers to identify and adopt a single pair of detection and quantitation limit approaches that can meet CWA needs, EPA also believes that any approach should be acceptable for use if it meets all of the criteria established above and fulfills the needs of the specific CWA application in which it should be used.

The Clean Water Act authorizes state or local governments to implement specific aspects of the Act, with the proviso that they do so in a way that is at least as protective (i.e., stringent) as the national

standards put forth by EPA. Therefore, this criterion is intended to ensure that any detection and quantitation limit approach adopted by the Office of Water is sufficiently clear and defined that it allows for comparison with approaches adopted by state or local governments. It is important to note that this criterion does not establish the need for an approach or procedure that is less stringent than those already in use by state or local governments.

Finally, it is important to differentiate between detection and quantitation limit approaches and compliance evaluation thresholds. Detection and quantitation limit approaches pertain to measurement process thresholds. More specifically, a detection limit describes the lowest concentration at which it is possible to determine that a substance is present with some stated confidence, and a quantitation limit describes the lowest concentration at which it is possible to quantify the amount of a substance that is present. In contrast, compliance evaluation thresholds are used to support wastewater discharge limits established in National Pollutant Discharge Elimination System (NPDES) or pretreatment program permits. Such limits are usually expressed as either a maximum concentration of pollutant allowed in the discharge or a maximum mass of pollutant allowed to be discharged in a specific time period.

Ideally, analytical methods are available to allow for detection and quantitation of pollutants at concentrations that are lower than the discharge levels needed to protect or restore the quality of the receiving water. When such measurement capability does not exist, permitting authorities must decide how to incorporate detection and quantitation limits into the discharge permit. Historically, EPA has recommended that in such cases, the permitting authority include the water quality-based limit in the permit, but establish the compliance evaluation threshold at the quantitation limit of the most sensitive available method. However, as with other aspects of the Clean Water Act, state and local governments may adopt permitting and compliance evaluation approaches that are at least as stringent as those put forth by EPA, and some states have preferred to use the detection limit as the compliance evaluation threshold.

In examining each approach against this criterion EPA will consider 1) the applicability of various detection/quantitation approaches to the variety of data gathering decisions that must be made under the CWA, including those that do and those that do not involve compliance monitoring, and 2) the ability of the approaches to support state and local obligations for implementing the CWA.

Table 4-1. Relationship of Issues Considered in Chapter 3 to Evaluation Criteria Established in Chapter 4

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
All sections in Chapter 3	Any approach adopted by EPA must be scientifically valid. Although not explicitly discussed in Chapter 3, the need for scientific validity has been an underlying condition throughout EPA's assessment.	<i>Criterion 1:</i> The concept of scientific validity is widely accepted but loosely defined. For the purposes of establishing scientific validity in this evaluation, EPA has adopted conditions discussed by the U.S. Supreme Court as considerations pertaining to assessments of scientific validity when considering the admissibility of expert scientific testimony. These conditions are that it can be (and has been) tested; it has been subjected to peer review and publication; the error rate associated with the approach or methodology is either known or can be estimated; standards exist and can be maintained to control its operation; and it has attracted (i.e., achieved) widespread acceptance within a relevant scientific community.
3.2.2, Descriptive vs. Prescriptive Uses of Lower Limits to Measurement	In order to protect human health and the environment, EPA must measure pollutants at ever lower concentrations. Establishing stringent standards and a compliance scheme for laboratories is one way to more rapidly develop the ability to measure at these concentrations. A prescriptive strategy concerning detection and quantitation limits would be to: determine these limits at one or more well-operated laboratories; use the performance of these laboratories as the basis to establish limits for the method; and use the established limits as a performance standard that must be demonstrated by laboratories that practice the method. The use of such an approach is consistent with EPA's use of other prescriptive laboratory performance standards and would ensure that prescriptive detection and quantitation limits (i.e., performance standards) reflect the capabilities of a well-performing laboratory or laboratories. This is in contrast to a descriptive approach that would base performance on a population of laboratories that may not be representative of the best possible performance.	<i>Criterion 2:</i> "... laboratory and method performance can be affected by the use of performance criteria that serve as prediction or tolerance limits. Examples of such criteria include measures to demonstrate that a laboratory is producing accurate results at a concentration of interest..., measures to demonstrate that results are not biased by contamination..., and measures to demonstrate that the laboratory can achieve the sensitivity required to reliably detect pollutants at low concentrations (i.e., at the detection limit). It is likely that laboratory performance will be better (and variability will be lower) when laboratories are required to meet specified performance criteria in order to report results." <i>Criterion 4:</i> "In evaluating this criterion, EPA will favor procedures that reflect routine analytical conditions in a well-operated laboratory." <i>Criterion 5:</i> "In evaluating the approaches, EPA will give preference to those approaches that strike a reasonable balance between using state-of-the-art laboratories and a highly varied community of laboratories to establish quantitation limits."
3.3.1, Sources of Variability	There are a number of ways in which variability can be controlled. However, it is not possible to completely eliminate all variability within or between laboratories. Even if prescribed quality control and variability control procedures are in place, it should be recognized that some laboratories may achieve lower detection and quantitation limits than others. The potential effects of sources of variability should be considered when establishing detection and quantitation limit approaches.	<i>Criterion 2:</i> "... laboratory and method performance can be affected by the use of performance criteria that serve as prediction or tolerance levels... In examining each approach against this criterion, EPA will evaluate if the approach can be used to provide a realistic expectation of laboratory performance. As part of this assessment, EPA will examine the sources of variability captured by the approach, and the degree to which the statistics that underlie the approaches realistically reflect these sources of variability."

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.3.7, Statistical Prediction and Tolerance	<p>Percentiles and prediction and tolerance intervals are statistical tools for describing how something that already exists (percentiles) and describing a future occurrence (prediction and tolerance limits). Percentiles are fairly straight forward to interpret, i.e., they specify the percentage of a distribution that false below a given percentile value. Prediction and tolerance limits are, in effect, confidence limits on percentiles and can be somewhat more difficult to apply... Statistical intervals can, and have by a number of authors, be adapted for use in setting detection and quantitation levels... However, the use of prediction and/or tolerance limits in setting detection and quantitation limits is not an absolute requirement and should be evaluated in the context of specific applications and policy considerations. In practice, the effect of adjustment of detection and quantitation limits by use of prediction and tolerance intervals can be quite large, depending on the amount of available data and the choices of percentiles and confidence levels.</p>	
3.3.8, Design of Detection and Quantitation Studies	<p>Studies designed to characterize sensitivity can be affected by the selection of spiking concentrations in studies, how well uncontrollable factors in the measurement process are reduced, the degree to which the entire measurement process is studied, and the flexibility of the design factors in terms of the physical measurement. Resources may be insufficient to support detection/quantitation limit approaches that model variability versus concentration because the selection of concentrations may require iteration when results do not meet their respective criteria.</p>	<p>Criterion 2: "In examining this criterion, EPA will evaluate if the approach can be used to provide a realistic expectation of laboratory performance. As part of this assessment, EPA will examine the sources of variability captured by the approach, and the degree to which the statistics that underlie the approach realistically reflect these sources of variability."</p>

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.3.3, Outliers	<p>One or more statistical procedures may be used to identify extremely large or small measurement values (outliers). Because extreme values are expected to occur, it is not necessarily appropriate to exclude them from measurement results used to develop detection or quantitation values. Ideally, the analyst's records should be reviewed to establish if an extreme value was caused by failure to follow the method or by some rare event associated with the method. In large detection and quantitation studies, it may not be feasible to review all extreme values to determine if they are outliers. In such cases, removing all extreme values as if they were outliers may be acceptable, but study documentation should state this is the case and the percentage of data removed. Removing large percentages of extreme values may cause variability estimates to be understated, indicate that there are systematic problems with following the method, or indicate that there are problems with the procedure for determining the extreme values.</p>	<p><i>Criterion 2:</i> "A further consideration concerning routine variability is the means for rejection of outliers. True outliers can occur in laboratory data and some means of resolving outlier issues must be included. Statistical procedures are available for the identification of candidate outlier values. Once a candidate outlier has been identified, evaluation of the value from a chemical analytical perspective (e.g., some procedural error or quality control error has occurred) should be the basis of exclusion of the value from a data set. In cases where no cause for the outlier has been identified it may reasonable to reject an outlier on statistical grounds but every effort should be made to justify the exclusion on technical grounds."</p>
3.1.3, Matrix Effects	<p>Reference matrices should be used to establish method detection and quantitation limits. The procedures used to define detection and quantitation limits should allow for evaluation of data collected in particular matrices of concern. Matrix-specific determinations should be used only after all efforts to resolve matrix interferences have been exhausted.</p>	<p><i>Criterion 3:</i> "The reality of environmental analysis is that measurement capabilities generally improve over time. This is attributable to a number of factors... In each case, the improvements may not be observed across the entire laboratory community... In evaluating this criterion, EPA will favor affordable and easy-use procedures that allow analysts in a single laboratory to 1) determine matrix -specific variations based on real data and 2) demonstrate that lower detection and quantitation limit approaches associated with improvements in their measurement capabilities."</p>
3.1.4, Measurement Quality over the Life of a Method	<p>Given that measurement capabilities generally improve over time, EPA believes that detection and quantitation limit approaches should be supported by procedures that will allow individual laboratories and other organizations to affordably characterize such improvements.</p>	
3.2.1.2, Method Performance Verification by a Laboratory	<p>Even where a method describes the sensitivity measured or estimated by the developer or the organization that published the method, some means is needed to demonstrate that given laboratory can achieve sufficient sensitivity to satisfy the regulatory decision (e.g., monitoring compliance).</p>	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.2.6, Cost and Implementation Issues	The financial and technical resources required to determine detection limit approaches vary widely according to the complexity of the procedures involved. Organizations that develop methods typically have greater resources available for determining limits than do organizations that use the methods. EPA must be sensitive to the capabilities of the organizations that develop and use the methods. Data from EPA studies indicate that the true detection/quantitation limits can only be arrived at by running hundreds of replicates. A better alternative would be to identify a simple procedure that yields a reproducible estimate and to allow laboratory-specific adjustment based on actual conditions in the laboratory.	<i>Criterion 3:</i> "Any approach or procedure should be simple, complete, and cost effective to implement. The laboratories that can be expected to use detection/quantitation procedures will range from large laboratories and laboratory chains with a wide range of technical capability to "mom and pop" laboratories operated by one or few people with a limited set of statistical skills. If a procedure is complicated it will be error prone in its use... if a procedure requires investment of extensive resources... laboratories will have a disincentive to use the procedure. Therefore, if the Agency wishes to encourage the development and use of innovative techniques that improve measurement performance or lower measurement cost, the Agency must consider practicality and affordability as significant, if not co-equal, considerations to scientific validity."
3.3.4, Criteria for the Selection and Appropriate Use of Statistical Models	What can be sometimes overlooked in considering estimation for model fitting is that direct measurement of variation of the blank or low-level concentration may be the most cost-effective and least difficult method to implement. The loss in statistical efficiency in comparison to more elaborate estimation and model fitting methodology would be offset by the relative ease and lower cost.	
3.3.6, False Positives and False Negatives	A common error in many published discussions of false negatives in relation to detection and quantitation is the claim that using Currie's detection limit (as opposed to the critical level) as a reporting limit or action level will somehow "control" false negatives. That claim is both false and counter-productive... As long as the only tool for setting requirements for false positive and false negative measurement results is the reporting limit, setting the reporting limit higher reduces the probability of a false positive at the expense of increasing the probability of a false negative.	<i>Criterion 4:</i> "Any detection limit approach should be capable of providing regulators, the regulated community, and data users with confidence that a pollutant reported as being present really is present. Historically, nearly every detection approach has set the criterion for detection at 99 percent confidence... This criterion results in the probability of a false positive (i.e., that a pollutant will be stated as being present when it actually is not [a Type 1 error]) of one percent."
3.1.1, Blank vs. Zero Concentration	Useful detection and quantitation limit approaches should address the potential contribution of the blank, through both the design of the study that generates the detection and quantitation limit estimates and evaluation of study results.	<i>Criterion 4:</i> "In evaluating this criterion, EPA will favor procedures that reflect routine analytical conditions in a well-operated laboratory. For example, the procedure must be capable of arriving a detection limit when the substance of interest is not found in a blank and/or when instrument thresholds are adjusted for routine operation."
3.1.2, Lack of Instrument Response	Procedures for establishing detection or quantitation limits should take into account the impact of instrument non-response.	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.3.4, Criteria for the Selection and Appropriate Use of Statistical Models	<ul style="list-style-type: none"> Method sensitivity is usually established based on measurement variation. Nearly all analytical techniques produce results that can generally be classified according to one of three basic models. The LOQ advanced by Currie and ACS, and EPA's ML result from multiplying the standard deviation of replicate analyses by a factor of 10. This factor of 10 is directed at achieving a relative standard deviation of 10 percent. An advantage of this approach is that a quantitation limit is produced, regardless of what the RSD turns out to be. Another means of arriving at a limiting RSD is to graph RSD versus concentration. This approach is used by the ASTM International IQE. It has the advantage that a model is fit to data, rather than using a point estimate such as the LOQ or ML. However, it requires considerably more data than approaches based on point estimates, and how a model is selected can play a major role in the outcome. 	<p><i>Criterion 5:</i> "Measurement capabilities among laboratories vary depending on a number of factors, including, but not limited to, instrumentation, training, and experience. Similarly, measurement capabilities among different analytical methods vary depending on a number of factors, including the techniques and instrumentation employed and the clarity of the method itself... Historical approaches to recognizing method capabilities also have varied between those that allow the error expressed as relative standard deviation, or RSD among low-level measurements to vary, depending on the capabilities of the method, and those that fix this error (RSD) at a specific level."</p> <p>"EPA will evaluate various approaches against this criterion by examining the ease of adjustment of the RSD or other performance measure in the context of the measurement capability of the laboratory or the need to adjust measurement error to allow for environmental decisions. In evaluating the approaches, EPA will give preference to those approaches that strike a reasonable balance between using state-of-the-art laboratories and a highly varied community of laboratories to establish quantitation limits."</p>
3.2.1.3, NPDES	The NPDES system serves as the primary means by which EPA, states, and Tribes control point source releases into the nation's waters. Under this system individual facilities are issued NPDES permits that provide limitations on the type, concentration, and volume of pollutants that may be legally discharged. Typically, these pollutant controls reflect technology-based standards. If, however, these technology-based controls are not adequate to protect the water quality standard designated for the facility's receiving water, stricter controls are warranted. In such cases, NPDES permits contain water quality-based controls.	<p><i>Criterion 6:</i> "... it is important to differentiate between detection and quantitation limit approaches and compliance evaluation thresholds. Detection and quantitation limit approaches pertain to measurement process thresholds. More specifically, a detection limit describes the lowest concentration at which it is possible to determine that a substance is present with some stated confidence, and a quantitation limit describes the lowest concentration at which it is possible to quantify the amount of a substance that is present. In contrast, compliance evaluation thresholds are used to support wastewater discharge limits established in National Pollutant Discharge Elimination System (NPDES) or pretreatment program permits."</p>
3.2.3, Compliance Evaluation Thresholds	A situation that arises frequently in addressing water quality-based limits is the setting of the permit limit below the detection or quantitation limit of the most sensitive, approved analytical method. Permit writers should have the flexibility to use the detection limit, the quantitation limit, or other limit as the compliance evaluation threshold so that the environment is protected.	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.2.4, Accepting the Procedures of Voluntary Consensus Standards Bodies	<ul style="list-style-type: none"> The National Technology Transfer and Advancement Act (NTTAA) encourages Federal agencies to focus on increasing their use of voluntary consensus standards whenever possible, and gives Federal agencies discretion to use other standards where use of voluntary consensus standards would be inconsistent with applicable law or otherwise impractical. Two types of technical standards apply to NTTAA; a performance standard and a prescriptive standard. NTTAA does not direct agencies to favor one type of standard over another. One option is for EPA to employ a performance-based approach to establishing detection and quantitation limits, in which method developers, laboratories, and others would be free to use any one of a variety of approaches to establishing these limits, including the existing MDL procedure, or a VCSB. Thus, establishing method sensitivity would be considered a performance standard under NTTAA, rather than a prescriptive standard. The fact that different approaches (prescriptive standards) yield different answers would be immaterial if EPA evaluates the answers (e.g., the detection limit that is determined) relative to a specific decision (e.g., the regulatory limit for a given pollutant). 	<p><i>Criterion 6:</i> “The Clean Water Act requires EPA to conduct, implement, and oversee a variety of data gathering programs... In addition, EPA needs to apply detection to permitting, compliance monitoring, and other uses of the 40 CFR part 136 methods. These applications include: permitting; ambient and effluent compliance monitoring under NPDES and the pretreatment program; ambient and effluent compliance monitoring under state and local programs; quality control in analytical laboratories; and method promulgation...In theory, EPA could evaluate each of these applications independently and identify a detection and quantitation limit approach that is best suited to each application... EPA believes that such an approach would increase confusion, increase record keeping burdens, and increase laboratory testing burdens. For these reasons, EPA believes it is desirable to adopt a single pair of related detection and quantitation procedures that can be used to address all Clean Water Act applications... In examining each approach against this criterion, EPA will consider 1) the applicability of various detection/quantitation approaches to the variety of data gathering decisions that must be made under the CWA, including those that do and those that do not involve compliance monitoring, and 2) the ability of the approaches to support state and local obligations for implementing the CWA.”</p>
3.2.1.1, Method Development and Promulgation	<ul style="list-style-type: none"> EPA believes it would be impractical to force standardization on a single detection or quantitation limit approach on method developers and promulgate only those methods that contain the standardized approach. EPA also believes there are real benefits to standardization and that 1) all new methods developed by EPA for promulgation at 40 CFR part 136 should reflect such standardization, and 2) EPA should strongly encourage outside organizations to include these approaches in their methods. 	

Chapter 3 Section	Summary of Issue Discussion	Evaluation Criterion in Chapter 4
3.2.7, Use of a Pair of Related Detection and Quantitation Procedures	Although EPA could develop a separate detection and quantitation limit approach for each application and attempt to define and evaluate each of the separate approaches, the resulting matrix of approaches would cause confusion to regulators, permittees, and the laboratory community. Further, when proposed, each item in the matrix of approaches and applications would, individually, be subject to contention and second-guessing, and it is likely that the outcome would be nearly the same as if a single pair of approaches is selected. To avoid this outcome, EPA believes it is desirable to use a single pair of related detection and quantitation procedures to meet needs where they exist in all CWA applications.	<i>Criterion 6:</i> In theory, EPA could evaluate each of these applications independently and identify a detection and quantitation limit approach that is best suited to each application... EPA believes that such an approach would increase confusion, increase record keeping burdens, and increase laboratory testing burdens. For these reasons, EPA believes it is desirable to adopt a single pair of related detection and quantitation procedures that can be used to address all Clean Water Act applications.
3.2.5, National versus Local Standards for Measurement	CWA authorizes states and local governments to implement permits, with the requirement that they be at least as protective (stringent) as the national standards established by EPA. Thus, EPA must take into account the impact of any revised or new detection/quantitation limit approaches and procedures on state and local governments, as well as on those affected by state and local requirements.	<i>Criterion 6:</i> "This criterion will be evaluated by studying ... 2) the ability of the approaches to support state and local obligations for implementing the CWA."
3.3.2, Censoring Measurement Results	Measurement results are often reported as less than some detection, quantitation, or reporting limit (i.e., they are censored below a designated limit). The primary reason for censoring is to avoid reporting highly unreliable results. Although such results may have high measurement error in a relative sense, they are of value to statisticians and modelers who are interested in analysis and modeling of measurement processes.	None. Although the issue of censoring is important, it should not be a consideration when selecting a detection and quantitation limit approach. The decision to censor data is a data reporting and data use issue.

This chapter summarizes EPA's assessment of various detection and quantitation limit approaches against the evaluation criteria established in Chapter 4. Assessments of detection limit approaches are presented in Section 5.1 and include an assessment of the:

- EPA method detection limit (MDL; Section 5.1.1),
- ASTM International interlaboratory detection estimate (IDE; Section 5.1.2),
- American Chemical Society (ACS) limit of detection (LOD; Section 5.1.3),
- International Organization for Standardization/International Union of Pure and Applied Chemistry (ISO/IUPAC) critical value (CRV; Section 5.1.4), and
- ISO/IUPAC minimum detectable value (MDV; Section 5.1.5).

Assessments of quantitation limit approaches are presented in Section 5.2 and include an assessment of the:

- EPA minimum level of quantitation (ML; Section 5.2.1),
- ASTM International interlaboratory quantitation estimate (IQE; Section 5.2.2),
- ACS limit of quantitation (LOQ; Section 5.2.3), and
- ISO/IUPAC LOQ (section 5.2.4).

A brief summary of the evaluation is presented in Tables 5-1 (detection limit approaches) and 5-2 (quantitation limit approaches).

EPA limited the assessment to detection and quantitation limit approaches advanced by ASTM International, ACS, ISO/IUPAC, and EPA, for use in EPA's Clean Water Act (CWA) programs, because these approaches are the most widely published and pertinent.

5.1 Detection Limit Approaches

Sections 5.1.1 through 5.1.5 describe EPA's assessment of five detection limit approaches. Each discussion is divided into two major subsections. The first subsection describes the approach and, where applicable, the procedure that supports the approach. The second subsection details EPA's assessment of the approach based on the five criteria established in Chapter 4 for evaluating detection limit approaches.

Note: Six criteria are given in Chapter 4. Four of these pertain to both detection and quantitation limit approaches. Criterion 4 pertains only to detection limit approaches and Criterion 5 pertains only to quantitation limit approaches. Therefore, the discussions of each detection and quantitation limit approach that follow will omit the criterion that does not apply.

5.1.1 Evaluation of the MDL

Section 5.1.1.1 provides an overview of the MDL approach and the procedures used to implement the approach. Section 5.1.1.2 describes EPA's assessment of the MDL against the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 4 and 6).

5.1.1.1 Description of the MDL Approach and Procedure

As promulgated at 40 CFR part 136, Appendix B, the MDL is defined as:

“the minimum concentration of a substance that can be measured and reported with 99% confidence that the analyte concentration is greater than zero and is determined from analysis of a sample in a given matrix containing the analyte.”

A six-step procedure is given in Appendix B, with an optional seventh step to verify the reasonableness of the MDL determined in the first six steps. The procedure is intended for use by experienced analytical chemists. A brief summary of the MDL procedure is as follows:

1. The analyst makes an estimate of the detection limit based on one of four options: the instrument signal to noise ratio; three times the standard deviation of replicate blank measurements; a break in the slope of an instrument calibration curve; or known instrument limitations.
2. The analyst prepares a volume of reagent water that is as free of the target analyte as possible (if the MDL is to be determined in reagent water).
3. The analyst prepares a sufficient volume of spiked reagent water (or of an alternate matrix) to yield seven replicate aliquots that have a concentration of the target analyte that is at least equal to or in the same concentration range as the estimated detection limit (it is recommended that the concentration of the replicate aliquots be between 1 and 5 times the estimated detection limit).
4. All of the replicate aliquots are processed through the entire analytical method.
5. The variance (S^2) and standard deviation (S) of the replicate measurements are determined, as follows:

$$S^2 = \frac{1}{n - 1} \left[\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i \right)^2}{n} \right]$$

$$S = \sqrt{S^2}$$

where:

X_i ; $i=1$ to n , = are the analytical results in the final method reporting units obtained from the n sample aliquots and Σ refers to the sum of the X values from $i=1$ to n .

6. The MDL is then determined by multiplying the standard deviation (S) by the Student's t -statistic at a 99% percentile for $n-1$ degrees of freedom. If seven replicates are used, the Student's t -value is 3.143. This information is used to calculate the MDL as follows:

$$MDL = t_{(n-1, 1-\alpha = 0.99)} (S)$$

where:

MDL = the method detection limit

$t_{(n-1, 1-\alpha=.99)}$ = the Student's t -value appropriate for a 99% confidence level with $n-1$ degrees of freedom, and

S = the standard deviation of the replicate analyses.

A 95% confidence interval for the determined MDL may be calculated from percentiles of the chi square over degrees of freedom distribution (χ^2/df).

7. The optional iterative procedure to verify the reasonableness of the MDL involves spiking the matrix at the MDL that was determined in Step 6, and analyzing another seven replicates spiked at this level. The F-ratio of the variances (S^2) is determined and compared with the F-ratio found in the table, which is 3.05. If $S_A^2/S_B^2 > 3.05$, the analyst is instructed to respike at the most recently calculated MDL and process the samples through the procedure starting with Step 4. If $S_A^2/S_B^2 > 3.05$, then the pooled standard deviation is determined. The pooled standard deviation is then used to calculate the final MDL as follows:

$$MDL = 2.681 \times S_{pooled}$$

where 2.681 is equal to $t_{(12, 1-\alpha=.99)}$.

The 95% confidence interval around the final MDL may be determined using the chi squared over degrees of freedom distribution.

The MDL procedure given at 40 CFR part 136, Appendix B is described as being applicable to 1) a wide variety of sample types, ranging from reagent water containing the analyte of interest to wastewater containing the analyte of interest, and 2) a broad variety of physical and chemical measurements. To accomplish this, the procedure was made device- and instrument-independent.

5.1.1.2 Assessment of the MDL Against the Evaluation Criteria

The following five subsections discuss the MDL approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.1.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

For the purposes of evaluating scientific validity, EPA is using the conditions discussed by the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals* (1993) and *Kumho Tire Co. v. Carmichael*, (1999) (see Chapter 4, Criterion 1).

Condition 1: It can be (and has been) tested. The MDL procedure meets this condition. The MDL has been used experimentally since 1980 and in a regulatory context since 1984. The MDL procedure is the most widely used and, therefore, the most widely tested detection limit procedure in the history of approaches of detection.

Critics of the MDL have noted that the detection limit produced with the MDL procedure can vary depending on the spike levels used. This would suggest, on the surface, that the MDL procedure can be used to obtain results that do not support the MDL approach. This is a misinterpretation of the MDL

based on the mistaken assumption that spike levels may be arbitrarily selected. In fact, step 1) of the MDL procedure specifies a number of criteria, based on chemical analytical considerations, that must be met in selecting the spike levels (see Section 5.1.1.1, Step 1).

In preparation for the assessment of detection and quantitation approaches, EPA exhaustively tested the MDL procedure with 10 different techniques, at decreasing spike concentrations, to evaluate this concern and determine how well the procedure characterized the region of interest. Results of the study suggest that, although the calculated MDL could vary depending on the spike level used, the procedure was capable of reasonably estimating a detection limit when the full iterative procedure was employed. Given these findings, and the previously noted concern that acceptable spike levels have been subject to misunderstanding, EPA believes that Step 1 of the MDL procedure should be revised to improve reader understanding of appropriate spiking levels, and that the iterative procedure in Step 7 of the MDL procedure should be made mandatory for development or revision of an MDL published in an analytical method.

Condition 2: It has been subjected to peer review and publication. The MDL meets this condition. Prior to promulgation by EPA, the MDL approach and supporting procedure was published by Glaser *et al.* in a peer-reviewed journal (Glaser, *et al.*, 1981).

Condition 3: The error rate associated with the procedure is either known or can be estimated. It is possible to estimate error rates associated with the MDL procedure. It is also possible to calculate confidence intervals about estimated MDLs that are expressions of uncertainty in the estimates. Clarification is in order because the promulgated MDL definition may be somewhat confusing in some respects. In particular, the definition is confusing with regard to whether the MDL is a true concentration or a value estimated from measured data. Another source of confusion lies in terminology. Because the MDL employs the term “detection” and is based on the approaches developed by Currie, it has often been incorrectly assumed to be the equivalent of Currie’s “detection limit,” when in fact, it is the equivalent of Currie’s “critical value,” which is the point at which the detection decision is made. EPA believes that the approach of MDL can be clarified by slightly revising the definition as follows:

“The method detection limit (MDL) is an estimate of the measured concentration at which there is 99% confidence that a given analyte is present in a given sample matrix. The MDL is the concentration at which a decision is made regarding whether an analyte is detected by a given analytical method. The MDL is calculated from replicate analyses of a matrix containing the analyte and is functionally analogous to the “critical value” described by Currie (1968, 1995) and the Limit of Detection (LOD) described by the American Chemical Society (MacDougall, et al. 1980, and Keith, et al. 1983).”

Condition 4: Standards exist and can be maintained to control its operation. The MDL approach is supported by a clearly defined, published procedure to control its operation. The procedure gives the steps to be followed and instructs the analyst to use the entire measurement process. Hundreds, if not thousands, of laboratories have successfully implemented the MDL procedure since its promulgation in 1984. EPA has found that when laboratories are required to perform MDL studies as part of an interlaboratory study, the results reported by the laboratories are generally consistent (i.e., within the expected variability). EPA has observed similar consistency in use of the MDL by laboratories required to perform the procedure to demonstrate proficiency with a method. Therefore, the MDL meets this condition.

That said, however, EPA believes that additional guidance can be provided to clarify certain aspects of the MDL procedure, particularly with respect to handling outliers, the optional reasonableness step, and multi-analyte test methods. The MDL procedure contains no discussion of outliers. It may be

helpful to clarify that 1) results should be discarded only if the results are associated with a known error that occurred during analysis (e.g., the replicate was spiked twice) or through a statistically accepted analysis of outliers, and 2) that laboratories should not run more than seven replicates and simply pick the best of the seven results. The optional step involves iterative testing to verify that the determined MDL is reasonable; EPA has observed that few organizations bother to perform this step. EPA also has observed that when a method involves a large number of analytes, it can be difficult to get all analytes to pass the iterative test in the same run. The MDL procedure would benefit from the addition of guidance on how and when to address each of these issues.

In addition, EPA notes that the calculation of the 95% confidence interval described in Step 7 is neither routinely performed by laboratories, nor are the results employed by regulatory agencies, including EPA. Therefore, EPA believes that the MDL procedure could be streamlined by deleting this calculation.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The MDL meets this condition. Within EPA, the MDL has been used by the Office of Research and Development, Office of Science and Technology, Office of Ground Water and Drinking Water, Office of Solid Waste, Office of Emergency and Remedial Response, and other offices. The MDL also has been used outside of EPA in methods published by ASTM International, in *Standard Methods for the Examination of Water and Wastewater*, jointly published by the American Public Health Association (APHA), the American Water Works Association (AWWA), and the Water Environment Federation (WEF), and in methods elsewhere. Although the MDL has been criticized by some, EPA believes that it is the most widely used approach of detection within the environmental chemistry community. Many states incorporate the MDL into NPDES permits, for example, and laboratories often advertise MDLs in their sales literature.

5.1.1.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The MDL procedure is designed to demonstrate laboratory performance with a given method, and can be applied to a broad variety of physical and chemical methods. To accomplish this, the procedure was made device- or instrument- independent. The procedure also recognizes the importance of analyst experience, and explicitly directs the analyst to employ all sample processing and computation steps given in the analytical method when determining the MDL. (All of these aspects are addressed in the MDL procedure published at 40 CFR 136, Appendix B).

When the MDL procedure is followed as intended (i.e., the MDL is determined by an experienced analyst on each device or instrument used for a given method), the demonstrated MDL will include routine variability associated with the laboratory and the method. As noted in the previous section, EPA believes the MDL procedure could be improved by describing appropriate means for the identification and treatment of outliers. Such modifications would ensure that laboratories do not inappropriately discard replicate data when calculating MDLs.

EPA recognizes that one laboratory may obtain detection limits that are lower or higher than those in another laboratory. If the MDL is being determined during method development, it is important to determine the MDL at more than one laboratory to ensure the MDL published in the method reflects demonstrated expectations of method performance in a community of laboratories. EPA does not believe that this community should be so broad as to include the entire universe of possible laboratories that might desire to practice the method. Rather, EPA believes this community should include well-operated laboratories that are experienced with the techniques used in the method and that have some familiarity with the method.

In recent years, EPA's Office of Science and Technology has used single-laboratory studies to develop an initial estimate of the MDL for a new or modified method, and has verified these limits in interlaboratory studies or by conducting additional single-laboratory studies in other laboratories. For example, when EPA initially drafted Method 1631 for measurement of mercury, EPA estimated the MDL to be 0.05 ng/L based on results produced by a contract research laboratory. Additional single-laboratory MDL studies conducted in other laboratories suggested that the MDL should be raised to 0.2 ng/L to better reflect existing capabilities of the measurement community. During EPA's interlaboratory study, each laboratory was asked to conduct an MDL study. Every laboratory in the interlaboratory study met the MDL of 0.2 ng/L, the value published in the promulgated version of Method 1631.

EPA believes that 1) the MDL procedure does address demonstrated expectations of laboratory and method performance, including routine variability, and 2) if the MDL procedure is being employed for method development purposes, it should be performed in multiple laboratories to ensure that it adequately demonstrates expectations in a community of qualified laboratories.

5.1.1.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The MDL is designed for use by a single laboratory. The promulgated version of the MDL procedure can be performed with as few as seven analyses. If the MDL is to be determined in a matrix other than reagent water, additional analyses will be needed.

Use of the optional iterative procedure would increase the number of analyses by seven each time the procedure is implemented. If the procedure is implemented two times in reagent water, a total of 14 analyses are required. If the procedure is implemented two times in an alternative matrix, EPA estimates that 17-20 analyses may be required, given the possible need to determine the background concentration of the analyte in the alternative matrix. In any of these scenarios, the entire MDL determination can be performed in a single analytical batch (most EPA methods specify batch sizes of 20 samples). As a result, EPA believes that the MDL is among the most affordable of the procedures that have been suggested for determining detection limits. In terms of cost, the only approach that compares favorably with the MDL is the instrument detection limit (IDL). Although most versions of the IDL compare favorably in terms of the number of samples analyzed, the requirement to perform the test on three non-consecutive days has the potential to disrupt routine laboratory operations on three days instead of one. In addition, the IDL does not include sample preparation steps and, therefore, does not completely characterize a method.

5.1.1.2.4 Criterion 4: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.

EPA believes the MDL meets this condition and refers the reader to the discussion of this subject under Section 5.1.1.2.1, Condition 3.

5.1.1.2.5 Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.

The MDL meets this criterion. The MDL has been successfully applied to a variety of decisions under the CWA since 1984. In addition, many states and others have adopted the MDL in their own programs.

5.1.2 Evaluation of the ASTM International Interlaboratory Detection Estimate (IDE)

The interlaboratory detection estimate (IDE) was developed by ASTM International with support from members of the regulated industry in an attempt to provide a scientifically sound, comprehensive detection limit procedure that addresses the concerns of the regulated industry, of statisticians, and of analysts involved in ASTM Committee D 19 on water.

A brief summary of the procedure is given in Section 5.1.2.1 and Section 5.1.2.2 presents EPA's assessment of the IDE against the five criteria established for evaluating detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.2.1 Description of the IDE Approach and Procedure

ASTM Designation D 6091 is the *Standard Practice for 99 %/95 % Interlaboratory Detection Estimate (IDE) for Analytical Methods with Negligible Calibration Error*. As stated in the practice:

"The IDE is computed to be the lowest concentration at which there is 90 % confidence that a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have a true detection probability of at least 95 % and a true nondetection probability of at least 99 % (when measuring a blank sample)."

The IDE is determined and verified using a procedure containing 5 major steps with approximately 53 substeps and conditions. The full text of the IDE procedure is available from ASTM International. The five major steps and their functions are given in Section 6 of the IDE procedure and are as follows:

1. Overview of the procedure.
2. IDE Study Plan, Design, and Protocol - in this section, the task manager (study supervisor) chooses the analyte, matrix, and analytical method. Details are given for range finding; the concentrations to be used in the study; the study protocol (ASTM Practice D 2777 is suggested); the allowable sources of variation; and the number of laboratories, analysts, and days over which the study will be conducted.
3. Conduct the IDE Study, Screen the Data, and Choose a Model - after the study data are collected and screened according to ASTM Practice D 2777, interlaboratory standard deviation (ILSD) versus concentration data are tabulated and one of three models is fit to the data. The first attempt is at fitting a constant model. If the attempt fails, a straight-line model is attempted. If the straight-line model fails, an exponential model is fitted. After fitting, the model is evaluated for reasonableness and lack of fit. If the model fails, the study supervisor determines if a subset of the data should be analyzed or if more data are needed.
4. Compute the IDE - the IDE is computed using the ILSD model selected in Step 3 to estimate the interlaboratory standard deviation at a true concentration of zero and at the IDE, using a mean recovery model to transform measured and true concentrations. The IDE is computed as a one-sided 90 % confidence upper statistical tolerance limit.
5. Nontrivial Amount of Censored Data - this section addresses the effect of "non-detects" or "less-thans." Suggestions are given to see if uncensored data can be obtained from the laboratories or if the

study needs to be augmented with additional data. Suggestions are given for fitting a model to data that contain less than 10 % non-detects or less-thans to produce an IDE.

5.1.2.2 Assessment of the IDE Against the Evaluation Criteria

The following five subsections discuss the IDE approach and procedure in the context of the five evaluation criteria that concern detection limit approaches.

5.1.2.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. EPA is not aware of any organization, including ASTM International, that has conducted a study to test the procedure as written (i.e., designed and implemented an interlaboratory study that involves estimating an initial IDE [IDE₀] and multilaboratory analyses of multiple concentrations of each matrix of interest surrounding IDE₀). Developers of the approach performed limited testing of the approach on 1) simulated data sets and 2) real-world data sets generated for other purposes. However, these real-world data sets are of limited value for testing the IDE because the concentration ranges associated with the data are above the low-level region of interest. As part of this reassessment, EPA tested a variant of the IDE procedure on single-laboratory data sets designed for characterization of an analytical method in the region of detection. Despite the lack of comprehensive testing, EPA believes that the procedure can be tested, and therefore meets part of this condition. Specifically, the IDE meets the condition that it can be tested, but it only partially meets the condition that it has been tested.

Condition 2: It has been subjected to peer review and publication. Although the IDE has not been published in the peer-reviewed scientific literature, the IDE has undergone extensive review and ballot by members of ASTM Committee D 19, many of whom are qualified peer reviewers. Therefore, although the IDE does not meet this condition in the sense of formal peer review and publication, EPA believes it does meet the intent of this condition (i.e., submission to scrutiny of the scientific community). In addition, the IDE was reviewed by four peer reviewers as part of EPA's assessment of detection and quantitation limit approaches.

Condition 3: The error rate associated with the procedure is either known or can be estimated. In theory, expert statisticians could estimate the error rate of the IDE. However, the IDE procedure is extremely complex from an analytical chemistry and statistical perspective. As a result, it is unlikely that the error rate could be estimated by the typical users of the analytical method to which it would be applied, or even by the typical developers of an analytical method. Moreover, EPA found the model selection procedure to be highly subjective, a situation likely to yield different IDEs from the same data set, depending on the staff involved in performing the calculations. In practice, such conditions make it impossible to estimate the actual error associated with the IDE. Therefore, the IDE fails this condition.

One of the four peer reviewers charged with evaluating EPA's assessment of detection and quantitation limit approaches concurred with EPA's assessment of the IDE, specifically stating, "*I agree that the IDE procedure as outlined is so complex as to make simple determination of error rates associated with it untenable.*" (Piegorsch, 2002)

Condition 4: Standards exist and can be maintained to control its operation. The IDE approach and procedure is supported by a published procedure (standard) to control its operation. The procedure gives the steps to be followed in determining the IDE and instructs the study supervisor how to gather the data and compute an IDE.

However, there are several "gray areas" in the published procedure. The most significant gray area is in the description of model selection. The procedure provides insufficient guidance on use of residual plots to evaluate and select models and, as a result, selection of the model may be very subjective, especially if the number of concentrations is low. The discussion of what model to use after rejecting the exponential and linear model is also very vague. The Rocke and Lorenzato (hybrid) model is mentioned, as well as models with more than one coefficient. Much of the data evaluated by EPA have tended to suggest the exponential model, based on the statistical tests discussed. However, those data have almost always shown residual "patterns" when using this model, which would then lead to consideration of other models. In addition, fitting the constant model is never discussed in detail. Most likely, this is done by simply calculating a mean (weighted if necessary) of the variances from the different concentrations; however, such calculations are never explicitly stated.

Another concern with the standard is that it gives procedures that are inconsistent with procedures given in the IQE standard, even though the two approaches should be consistent for a given analyte with a given method. For example, the exponential model figures prominently in the IDE procedure, where it is one of the three main models discussed. The Rocke and Lorenzato model is not discussed in the IDE procedure, but it figures prominently in the IQE procedure. In theory, a single model should support the definition of both the detection and quantitation limits for a given analyte by a given method. As another example, the IDE procedure includes a multiplier to account for bias in estimating the true standard deviation with the sample standard deviation, but the IQE does not.

Finally, the procedure contains statistical errors that, if followed as written, could produce inaccurate IDE values. For example, Table 1 of the procedure contains "Computations to Estimate Straight-Line Model Coefficients by Means of Least Squares— Ordinary and Weighted," but the weighted least squares formulae given in the table are incorrect. The formulae for the weighted means of the spike values and results given in Table 1 of D6091 would only be appropriate if the weighting were done based on the number of replicates per spike level, rather than on the estimated variance calculated using the chosen standard deviation model.

In conclusion, EPA believes that although the IDE is supported by a published procedure, that procedure will not control its operation because of the degree of subjectivity involved implementing the procedure, errors in the procedure, and inconsistencies with its IQE counterpart. Therefore, the IDE fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IDE fails this condition because it is only familiar to, and has been accepted by, a very narrow segment of the scientific community. Although the IDE has been approved by ASTM for more than 5 years, EPA is not aware of an IDE that has been published in the open literature or in an analytical method, including an ASTM method.

5.1.2.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The IDE procedure is designed to reflect expectations of interlaboratory performance, including routine variability. The procedure contains extensive instructions for dealing with unusual conditions, including sources of variability and outliers. However, EPA studies of a single-laboratory variant of the procedure suggested that the procedure may not always work as intended. For example, model selection based upon hypothesis tests (as described in D6091, Section 6.3.3.2) almost always indicated that the exponential model should be used, even when the data seemed to be show constant or approximately linear error, while examination of residual plot indicated "systematic behavior" (i.e., non-random deviations from the model) for the exponential and linear models. Another concern with the IDE

procedure is that use of the non-mandatory appendices in ASTM D 6512 to determine the fit of a model may produce results that differ from those that would be obtained by using the default procedures for testing model fit that are built into off-the-shelf statistical software. Such observations, along with the concerns described in Section 5.1.2.2.1, condition 4, lead EPA to believe that, while the IDE approach addresses demonstrated expectations of laboratory and method performance, the IDE procedure does not adequately do so. Therefore, the IDE only partially meets this criterion.

5.1.2.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The IDE procedure is designed for use by an ASTM International study supervisor or task manager and not as a procedure that a single laboratory can use to evaluate method performance. EPA is aware that ASTM Committee D 19 is developing a Within-laboratory Detection Estimate (WDE), but the WDE is presently only in the formative stages. The WDE may meet this criterion, but the IDE does not.

Regarding cost, the IDE procedure would be the most costly of the procedures that EPA has evaluated because of the time it would take to understand and implement the procedure, and requirements for: 1) estimation of IDE₀, 2) interlaboratory data, 3) extensive statistical intervention in determining the correct model, and 4) possible reanalyses if the resulting IDE does not meet the criteria in the procedure.

5.1.2.2.4 Criterion 4: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.

By definition, the IDE is designed to achieve "a true detection probability of at least 95 % and a true nondetection probability of at least 99 %." Although the 99% probability of a "true nondetection" is equivalent to the 99% confidence that the substance is actually present given in Criterion 4, ASTM International also included the simultaneous requirement for a 95% probability of a "true detection." The developers are using the IDE as a means to control the rates of both false positive and false negative results, in essence, making the IDE analogous by definition and formulaic construction to the *detection limit* (DL) defined by Currie (1968). The IDE accomplishes this goal by using a tolerance limit that increases the IDE well above the point at which the detection decision would be made. For a discussion of this issue, see Sections 3.3.6 (false positives and false negatives) and 3.3.7 (prediction and tolerance intervals) in Chapter 3 of this document.

As noted in Section 2.1 of Chapter 2 of this document, Currie (1968) used the term *detection limit* (subsequently termed the *minimum detectable value*) to refer to a true concentration that has a high probability of generating measured values greater than the critical value. That is, measurements on samples that contain concentrations equal to the *detection limit* have a high probability of exceeding the *critical value* and are, therefore, unlikely to result in a decision that the substance is not detected in the sample. However, the *detection decision* is made on the basis of comparing sample measurements to the *critical value*. With regard to his definition of the "*detection limit*," Currie (1995) states "*The single, most important application of the detection limit is for planning.*"

When the allowance for false negatives and the prediction and tolerance limits are taken into account, the resulting IDE is raised to the point at which the probability of a false positive is less than 0.00000001 (10^{-8}). This protection against false positive results is excessive and would yield numerical values of little practical value for making the detection decision.

5.1.2.2.5 *Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.*

EPA's comparison of detection limits produced by various detection limit approaches shows that the median IDE is considerably higher than ACS, ISO/IUPAC, and EPA detection limits. Although the IDE could be applied to some decisions to be made under CWA, it may not support decisions when pollutant levels need to be protective of human health and the environment because the IDE is an implementation of Currie *detection level* or *minimum detectable value*, and may be considerably higher than these levels. At best, the IDE only partially meets this criterion.

5.1.3 Evaluation of the ACS Limit of Detection

The limit of detection (LOD) was developed by the Committee on Environmental Improvement (CEI) of the American Chemical Society (ACS). ACS is a professional society for chemists and other scientists and the publisher of a number of scientific journals. It is not a voluntary consensus standards body (VCSB), nor does it develop or publish analytical methods. In 1978, the ACS/CEI began addressing concerns about the lack of useful standards for interlaboratory comparisons. In 1980, the Committee published its *"Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry"* (MacDougall, *et al.*, 1980), which included the approaches of the LOD and the limit of quantitation (LOQ).

5.1.3.1 Description of the ACS LOD

The 1980 *"Guidelines"* define the LOD as:

"... the lowest concentration of an analyte that the analytical process can reliably detect. ... The LOD in most instrumental methods is based on the relationship between the gross analyte signal S_t , the field blank S_b , and the variability in the field blank σ_b ."

and construct the formal relations using the equation:

$$S_t - S_b \geq K_d \sigma$$

where K_d is a constant. ACS recommended a minimal value of 3 for K_d . Thus, the LOD is 3σ above the gross blank signal, S_b . In the 1980 publication, the ACS stated that at $K_d = 3$, there is a 7% risk of false negatives and false positives. Given that the LOD is 3σ above the blank, however, EPA believes that the risk of false positives is somewhat less than 1%.

In 1983, the ACS Committee published *"Principles of Environmental Analysis"* (Keith *et al.*, 1983). That publication occurred after the 1981 paper on the Method Detection Limit (MDL), and ACS/CEI stated that the LOD is numerically equivalent to the MDL as S_b approaches zero. However, neither the 1980 nor 1983 ACS publications provide a specific procedure for estimating the LOD, nor do they provide a minimum number of observations needed to estimate the gross blank signal or the variability term σ_b .

5.1.3.2 Assessment of the LOD Against the Evaluation Criteria

The following five subsections discuss the LOD approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.3.2.1 *Criterion 1: The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. Testing of the ACS LOD is hampered by 1) the lack of a supporting procedure for establishing an LOD, and 2) it's conceptual dependence on the variability associated with measuring blanks. For example, there is no procedure to govern the minimum number of analyses needed to characterize the variability of a blank sample. Because many environmental chemistry techniques yield a zero, or possibly even negative, value when a blank sample is analyzed, and because the LOD approach is based on the standard deviation of these results, directly testing the LOD in such techniques will yield a zero or negative value. One solution for testing is to rely on ACS' 1983 statement that the LOD is conceptually equivalent to the MDL as the blank signal approaches zero, and employ the MDL procedure as a means for indirectly testing the LOD approach. EPA believes that use of the MDL procedure is a viable means for testing the approach; therefore, the LOD meets this condition.

Condition 2: It has been subjected to peer review and publication. The LOD definition was published in the peer-reviewed journal *Analytical Chemistry* in 1980 and 1983. Therefore, the LOD meets this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The error rates can be estimated, so the LOD meets this condition. The error rate for both false positives and false negatives is stated to be 7 % in the 1980 *Analytical Chemistry* article. However, EPA believes that, because the LOD is stated to be 3 times the standard deviation of replicate measurements of a blank, the false positive rate is overstated and is actually somewhat less than 1 % whereas the false negative rate depends on the true concentration in the sample.

Condition 4: Standards exist and can be maintained to control its operation. The LOD lacks a clearly defined procedure for estimating the important terms required to derive it. Although it may be possible to derive LOD values from data used to derive EPA MDL values, there is no procedure giving explicit instructions on the use of replicate blanks, replicate spiked samples, or a minimum recommendation for the number of replicates. Therefore, the LOD fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because ACS does not develop and publish analytical methods, it is difficult to determine the degree of acceptance of the LOD. EPA has not specifically investigated the numbers of papers published in ACS journals that include LOD values, and EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the LOD in particular. However, ACS LOD values have appeared in the technical literature. Given that ACS is a relevant scientific community, and that use of the LOD has appeared in the technical literature, EPA believes the LOD meets this condition.

5.1.3.2.2 *Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.*

The LOD approach is designed to address demonstrated expectations of laboratory and method performance, including routine variability, and thus appears to meet this criterion. Unfortunately, ACS has not published a procedure to implement the approach. In other words, the LOD addresses demonstrated expectations of laboratory and method performance in theory, but in practice, provides no direct means for performing these demonstrations. Therefore, EPA believes the ACS LOD only partially meets this criterion.

5.1.3.2.3 *Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The ACS LOD approach is not supported by a clearly defined procedure for establishing the LOD. Therefore, it fails this criterion.

5.1.3.2.4 *Criterion 4: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

The 1983 publication associated the LOD with the "99% confidence level when the difference ($S_t - S_b$) $> 3\sigma$." Therefore, the LOD satisfies this criterion.

5.1.3.2.5 *Criterion 6: The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

In the absence of a procedure for determining LOD values, the ACS LOD fails to meet this criterion because it cannot be used in a regulatory context. The LOD passes only if it is assumed to be functionally equivalent to the MDL (i.e., the MDL procedure is used to establish an LOD).

5.1.4 Evaluation of the IUPAC/ISO Critical Value (CRV)

The critical value (CRV) was developed by the International Union of Pure and Applied Chemistry (IUPAC) and the International Organization for Standardization (ISO). IUPAC and ISO are professional societies for chemists and other scientists. ISO develops and publishes analytical methods through its Task Groups. In 1995, Lloyd Currie of the National Institute for Standards and Technology (NIST; formerly the National Bureau of Standards) published a signature discussion of IUPAC approaches for detection and quantitation (*Pure and Appl. Chem.* 67:10, 1699-1722). Although refined during the intervening years (see Currie, L.A., *J. Radiochem. And Nuclear Chem.* 245:1, 145-156, 2000), the CRV approach remains basically as described in 1995.

5.1.4.1 Description of the ISO/IUPAC Critical Value (CRV) Approach and Procedure

The 1995 article states that the critical value (L_c) is:

"... the minimum significant value of an estimated net signal or concentration, applied as a discriminator against background noise. This corresponds to a 1-sided significance test."

For a normal distribution with known variance, L_c reduces to:

$$L_c = z_{(1-\alpha)}\sigma_0$$

where:

$1-\alpha$ is the false positive error rate, recommended at 5 % ($\alpha = 0.05$), and σ_0 is the standard deviation at zero concentration

If σ_0 is estimated by s_0 (replicate measurements of a blank), $z_{(1-\alpha)}$ is replaced by the Student's t -value. For 7 replicates (6 degrees of freedom), the Student's t -value is 1.943, where $\alpha = 0.05$.

5.1.4.2 Assessment of the CRV Against the Evaluation Criteria

The following five subsections discuss the CRV approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.4.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. The lack of a supporting procedure for establishing the CRV, coupled with its conceptual dependence on the variability of blank measurements makes testing of the approach difficult. For example, if blank measurements fail to produce a response, it is impossible to calculate a CRV because the standard deviation of zero is zero. One solution for testing the approach is to assume that the CRV is functionally equivalent to the MDL as the blank signal approaches zero, and use a slightly modified version of the MDL procedure to test the CRV approach. The slight modification involves selecting a Student's t -value based on $\alpha = 0.05$ instead of $\alpha = 0.01$, for $n-1$ degrees of freedom. EPA believes this is a reasonable assumption, and therefore, that the MDL procedure is a viable means for testing the CRV approach. Therefore, the CRV meets this condition.

Condition 2: It has been subjected to peer review and publication. The IUPAC/ISO definitions meet this criterion. Moreover, it is likely that these definitions have received greater peer review than any of the other approaches.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The error rate is specified by α , with a suggested value of 0.05 (5%). Therefore, the CRV meets this condition.

Condition 4: Standards exist and can be maintained to control its operation. The CRV is defined in the various publications by Currie. However, EPA's search of the literature and the ISO web site found no standard for control of the approach. Therefore, the CRV fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because IUPAC and ISO are international bodies, it is difficult to determine the degree of acceptance of the CRV in the U.S. and the world community. EPA has not specifically investigated the number of papers in published journals that include CRV values, but EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the CRV in particular. Therefore, it is difficult to determine if the CRV meets this condition.

5.1.4.2.2 Criterion 2: *The approach should address demonstrated expectations of laboratory and method performance, including routine variability.*

The CRV approach is designed to account for the variability of measurements of the blank in the context of a "chemical measurement process" (method). Unfortunately, neither ISO, IUPAC, nor Currie have published a procedure to implement the approach. As a result, the CRV addresses demonstrated expectations of laboratory and method performance in theory, but in practice, provides no direct means for performing these demonstrations. Therefore, EPA believes the CRV partially meets this criterion.

5.1.4.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance*

The CRV approach is not supported by a clearly defined procedure for establishing a CRV. Therefore, the CRV fails this criterion.

5.1.4.2.4 Criterion 4: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

Although the CRV suggests $\alpha = 0.05$, resulting in $1-\alpha$ of 0.95 or 95 % probability of detection, the approach allows for the specification of other probabilities. Therefore, the CRV satisfies this criterion.

5.1.4.2.5 Criterion 6: *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.*

In the absence of a procedure for establishing CRVs, the CRV approach fails to meet this criterion because it cannot be used in a regulatory context. The CRV passes only if it is assumed to be functionally equivalent to an MDL determined with α set at 0.05 instead of 0.01 (i.e., if the MDL procedure, with $\alpha = 0.05$, is used to establish a CRV).

5.1.5 Evaluation of the IUPAC/ISO Detection Limit

The detection limit or minimum detectable value (MDV) was developed by IUPAC/ISO and published in the same papers as the CRV (Section 5.1.4)

5.1.5.1 Description of the IUPAC/ISO Detection Limit Procedure

The 1995 publications define the minimum detectable value (detection limit) as follows:

"The Minimum Detectable Value (MDV) ... [is] ... the net signal (or concentration) of that value (L_D) for which the false negative error is β , given L_C (or α)." (see the CRV for L_C)

For a normal distribution with known variance, L_D reduces to:

$$L_D = z_{(1-\beta)} \sigma_D$$

where:

z is the state variable

$1-\beta$ is the false negative error rate, recommended at 5 % ($\beta = 0.05$), and

σ_D is the standard deviation at the detection limit

Earlier publications refer to the minimum detectable value as the detection limit. To avoid confusion in terminology and to help distinguish the ISO/IUPAC approach from the MDL, LOD, and CRV, EPA will refer to the ISO/IUPAC detection limit as the Minimum Detectable Value, abbreviated as MDV.

5.1.5.2 Assessment of the ISO/IUPAC MDV Against the Evaluation Criteria

The following five subsections discuss the ISO/IUPAC MDV approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-4, and Criterion 6).

5.1.5.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. The lack of a supporting procedure for establishing the MDV makes testing of the approach difficult. However, EPA believes that the MDV can be tested using data similar to those used to generate MDL values. Therefore, the MDV meets this condition.

Condition 2: It has been subjected to peer review and publication. The IUPAC/ISO definitions meet this condition; moreover, it is likely that this definition has received greater peer review than any of the other approaches.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The error rates are specified by α and β , both with suggested values of 0.05 (5 %). Therefore, the error rate is known.

Condition 4: Standards exist and can be maintained to control its operation. The MDV is defined in the various publications by Currie. However, EPA's search of the literature and the ISO web site found no standard for control of the approach. Therefore, the MDV fails this criterion.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because IUPAC and ISO are international bodies, it is difficult to determine the degree of acceptance of the MDV in the U.S. and the world community. EPA has not specifically investigated the number of papers in published journals that include MDV values, but EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the MDV in particular. Therefore, it is difficult to determine if the CRV meets this criterion.

5.1.5.2.2 Criterion 2: *The approach should address demonstrated expectations of laboratory and method performance, including routine variability.*

The MDV approach is designed to account for the variability of measurements of the blank in the context of a "chemical measurement process" in the sense that it is used in concert with a critical value that is based on blank measurement variability. The MDV is the true concentration that is used in the planning of method evaluation and development. The actual detection decision is made at the critical value (CRV) which is determined from measured values. The approach of a true concentration MDV and its associated allowance for false negatives is of little practical value in making the actual detection decision. Therefore, the MDV fails this criterion. The allowance for false negatives in a regulatory context is discussed in greater detail in Chapter 3.

5.1.5.2.3 Criterion 3: *The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance*

The MDV approach is not supported by a clearly defined procedure for establishing MDV values. Therefore, the MDV fails this criterion.

5.1.5.2.4 Criterion 4: *The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.*

The allowance for false negatives raises the probability of detection to a value estimated to be greater than 99.999999 % (probability of a false positive less than 10^{-8}). This protection against false positive results is excessive and would yield numerical values of little practical value for making the detection decision. Perhaps more importantly, as noted by Currie (1995) and discussed in Section 5.1.2.2.4 of this document, the *detection decision* is made on the basis of comparing sample measurements to the *critical value*. Therefore, the MDV fails this criterion.

5.1.5.2.5 Criterion 6: *Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government*

In the absence of a procedure for establishing MDV values, the MDV approach fails to meet this criterion because it cannot be used in a regulatory context.

5.2 Quantitation Limit Approaches

Sections 5.2.1 through 5.2.4 describe EPA's assessment of four quantitation limit approaches. Each discussion is divided into two major subsections. The first subsection describes the approach and, where applicable, the procedure that supports the approach, and the second subsection details EPA's assessment of the approach based on the five criteria established in Chapter 4 for evaluating quantitation limit approaches.

Note: Six criteria are given in Chapter 4. Four of these pertain to both detection and quantitation limit approaches. Criterion 4 pertains only to detection limit approaches and Criterion 5 pertains only to quantitation limit approaches. Therefore, the discussions of each detection and quantitation limit approach that follow will omit the criterion that does not apply.

5.2.1 Assessment of the EPA Minimum level of Quantitation (ML)

Section 5.2.2.1 provides an overview of the ML approach and the procedures used to implement the approach. Section 5.2.2.2 contains EPA's assessment of the ML against the five evaluation criteria that concern quantitation limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.1.1 Description of the ML Approach and Procedures

The present definition of the ML includes a statement of the approach and the procedures used to establish the ML. This definition states that the ML is:

“the lowest level at which the entire analytical system must give a recognizable signal and acceptable calibration point for the analyte. It is equivalent to the concentration of the lowest calibration standard, assuming that all method-specified sample weights, volumes, and clean up procedures have been employed. The ML is calculated by multiplying the MDL by 3.18 and rounding the results to the number nearest to (1, 2, or 5) $\times 10^n$, where n is an integer.”

The ML is designed to provide a practical embodiment of the quantification level proposed by Currie and adopted by IUPAC. It is functionally analogous to Currie's "determination limit" (described in Chapter 2, Section 2.1) and the American Chemical Society's Limit of Quantitation (LOQ). The LOQ is discussed in Section 5.2.3 of this chapter. Chapter 2 (Section 2.2.2) describes the ML approach in additional detail.

The first part of the ML definition (i.e., the lowest level at which the system gives a recognizable signal and acceptable calibration point for the analyte) ties the quantification limit to the capabilities of the measurement system. The second part of the ML definition provides a procedural means for establishing the ML.

The procedural component of the definition is designed to yield an ML value that equals approximately 10 times the standard deviation of replicate analyses used to determine the MDL. (The exact value corresponding to 10 times the standard deviation is rounded to avoid error that would arise from preparation of calibration standards at exact, unrounded concentrations.) The procedure given in the above definition assumes that exactly seven replicates are used to determine the MDL. EPA has observed, however, that laboratories occasionally perform MDL studies with more than the required minimum of seven replicates. When this is done, the Student's *t*-value used to calculate the MDL should be adjusted accordingly. Similarly, the Student's *t*-value would need to be adjusted when a laboratory performs the optional iterative test described in Step 7 of the MDL procedure, or if outlier testing results in the use of less than seven replicates to establish the MDL. Therefore, EPA believes that the ML definition should be revised to eliminate the assumption of seven replicates and clarify its functional equivalence to Currie's critical value and ACS' LOQ. In addition, a detailed procedure should be developed to ensure proper calculation of the ML when more than seven replicates are used to establish the MDL or when iterative testing is used to establish the MDL.

5.2.1.2 Assessment of the ML against the Evaluation Criteria

The following five subsections discuss the ML approach and procedure in the context of the five evaluation criteria that concern quantitation limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.1.2.1 Criterion 1: The detection and quantitation limit approaches should be scientifically valid.

Condition 1: It can be (and has been) tested. The ML meets this condition. The ML has been used experimentally since 1979 and in the regulatory context since 1984. The ML is tested each time a laboratory calibrates an instrument because methods that include the ML require that it be included as the lowest non-zero standard in these calibrations.

Moreover, EPA exhaustively tested the MDL and ML procedure with 10 different techniques at decreasing spike concentrations to evaluate how well the MDL and ML procedures characterized the region of interest in preparation for this reassessment of detection and quantitation limit approaches. Results of the study suggest that 1) although the calculated MDL and ML could vary depending on the spike level used, the procedure was capable of reasonably estimating detection and quantitation limits when the full iterative MDL procedure was employed, and 2) the rounding process employed to determine the ML generally yielded consistent MLs even with slight variations in the calculated MDL.

In other words, if the procedure for establishing an ML is properly implemented for a given method, it will yield an ML value that is consistent with the approach, and this ML value will be verified (tested) by a laboratory each and every time it calibrates the instrument used to analyze samples by the method.

Condition 2: It has been subjected to peer review and publication. The ML has not been published in a peer reviewed journal. However, it was evaluated by four peer reviewers as part of EPA's assessment of detection and quantitation limits. These reviewers noted that:

"The MDL and ML concepts evaluated in Section 5.1.1 and 5.2.1, respectively, are shown in this evaluation to be technically sound and practical." (Wait, 2002)

"With respect to the limit of quantitation concept, the EPA ML is as good as any of the others given..." (Rocke, 2002)

"The MDL and ML have stood the test of time and provide a proven methodology which meets evaluation criteria stated in the TSD." (Cooke, 2002).

In addition, the present definition of the ML describes the approach and the procedures used to establish the ML. This definition is included in EPA Method 1631, which was extensively peer reviewed in accordance with EPA policies on peer review prior to publication and promulgation. Given that EPA's policies on peer review are as stringent as or more stringent than those used by many published journals, EPA believes that the ML has met a high standard of scientific review and scrutiny, and therefore, meets the intent of this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The uncertainty associated with any ML value can be calculated. EPA performed such calculations during this assessment and found that, on average across all techniques tested, the relative standard deviation of replicate measurements at the ML was approximately 7%. Median RSD values calculated for each multi-analyte method tested ranged from 6 to 14 percent. RSD values calculated for each single-analyte method tested ranged from 4 to 16 percent. (See Appendix C to this Assessment Document for a detailed discussion and presentation of results.)

Condition 4: Standards exist and can be maintained to control its operation. The ML meets this criterion. Detailed procedures (i.e., standards) for establishing the ML are given in the definition itself, although, as noted above, EPA believes that a detailed, stand-alone procedure should be created to ensure that the ML is properly calculated when other than seven replicates are used in its determination.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. EPA believes the ML meets this condition. The ML is functionally analogous to the American Chemical Society's LOQ and to the ISO/IUPAC quantification limit, suggesting widespread acceptance.

5.2.1.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The ML procedure is designed to provide a means by which a laboratory can demonstrate performance with a method under routine laboratory operating conditions. All recently developed EPA CWA methods require that a laboratory calibrate its instrument prior to analyzing environmental samples. The ML is defined as the lowest non-zero standard in the laboratory's calibration, and therefore, reflects realistic expectations of laboratory performance with a given method under routine laboratory conditions (i.e., under conditions of routine variability).

Also, the ML is based on the standard deviation of replicate analyses used to establish the MDL. As described in Section 5.1.1.2.2, these analyses are performed to characterize laboratory and method performance, including routine variability, at low concentrations. When a laboratory performs an MDL

study with seven replicates and multiplies the results by 3.18, the laboratory has demonstrated that it can achieve expected levels of performance at the ML.

EPA recognizes that one laboratory may obtain an MDL or ML that is lower or higher than those in another laboratory. If the ML is being established during method development, it is important to determine the ML at more than one laboratory to ensure that the published ML reflects demonstrated expectations of method performance in a community of laboratories. EPA does not believe this community should be so broad as to include the entire universe of possible laboratories that might desire to practice the method. Rather, EPA believes that this community should include well-operated laboratories that are experienced with the techniques used in the method and that have some familiarity with the method. See Section 5.1.1.2.2 for additional discussion of this topic.

5.2.1.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The ML is designed for use by a single laboratory. The ML can be directly determined from the MDL, which is among the most affordable of procedures for determining detection limits (see discussion in Section 5.1.1.2.3 for additional details regarding affordability). As a result, the ML is among the most affordable of procedures for determining quantitation limits.

5.2.1.2.4 Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.

The ML meets this criterion. The ML can be verified in a laboratory each time it calibrates an instrument. This calibration is dependent on identifying a recognizable signal for the analyte. In addition, because EPA includes the ML as the low point in the calibration range, that concentration is within the capabilities of the method, as demonstrated by either multiple single-laboratory studies or a multi-laboratory study of the method.

5.2.1.2.5 Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.

The ML meets this criterion. It has been used successfully to support state and local obligations under the Clean Water Act since 1984.

5.2.2 Assessment of the IQE

The Interlaboratory Quantitation Estimate (IQE) was developed by ASTM International with support from members of the regulated industry in an attempt to provide a scientifically sound, comprehensive quantitation limit procedure that addresses the concerns of the regulated industry, statisticians, and analysts involved in ASTM Committee D 19 on water. A brief summary of the procedure for establishing an IQE is given in Section 5.2.2.1. Section 5.2.2.2 presents EPA's assessment of the IQE against the five criteria established for evaluating quantitation limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.2.1 Description of the IQE Approach and Procedure

The ASTM Designation D 6512 is the *Standard Practice Interlaboratory Quantitation Estimate*. As stated in the practice:

"IQE_{Z%} is computed to be the lowest concentration for which a single measurement from a laboratory selected from the population of qualified laboratories represented in an interlaboratory study will have an estimated Z % relative standard deviation (Z % RSD, based on interlaboratory standard deviation), where Z is typically an integer multiple of 10, such as 10, 20, or 30, but Z can be less than 10."

The IQE is determined and verified using a procedure containing 5 major steps with approximately 46 substeps and conditions. The full text of the IQE procedure is available from ASTM International. The 5 major steps and their functions are given in Section 6 of the IQE procedure and are summarized below:

1. Overview of the procedure.
2. IQE Study Plan, Design, and Protocol - in this section, the task manager (study supervisor) chooses the analyte, matrix, and analytical method. Details are given for the appropriate range of study concentrations; the model of recovery vs. concentration; the study protocol (ASTM Practice D 2777 is suggested); the instructions to be given to the participating laboratories, including reporting requirements; the allowable sources of variation; and the number of laboratories, analysts, measurement systems, and days over which the study will be conducted.
3. Conduct the IQE Study, Screen the Data, and Choose a Model - after the study data are collected and screened according to ASTM Practice D 2777, the interlaboratory standard deviation (ILSD) versus concentration data are tabulated and one of three models is fit to the data. The first attempt is at fitting a constant model. If the attempt fails, a straight-line model is attempted. If the straight-line model fails, a hybrid (Rocke/Lorenzato) model is fit. After fitting, the model is evaluated for reasonableness and lack of fit. If the model fails, the study supervisor determines if a subset of the data should be analyzed or if more data are needed.
4. Compute the IQE - the IQE is computed using the ILSD model selected in Step 3 to estimate the relative standard deviation as a function of concentration. The first attempt is at 10% RSD (IQE_{10%}). If this attempt fails, IQE_{20%} is tried, then IQE_{30%}. IQEs greater than 30% are not recommended.
5. Nontrivial Amount of Censored Data - this section of the IQE procedure addresses the effect of "non-detects" or "less-thans." Suggestions are given to see if uncensored data can be obtained from the laboratories or if the study needs to be augmented with additional data. Suggestions are given for fitting a model to data that contain less than 10% non-detects or less-thans to produce an IQE.

5.2.2.2 Assessment of the IQE Against the Evaluation Criteria

The following five subsections discuss the IQE approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.2.2.1 *Criterion 1: The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. EPA is not aware of any organization, including ASTM, that has conducted a study to test the IQE procedure as written (i.e., designed and implemented an interlaboratory study involving multi-laboratory analysis of multiple concentrations of each matrix of interest). It has been tested by its developers using simulated data sets and on interlaboratory data sets that do not adequately characterize the low level region of interest. As part of this reassessment, EPA tested a variant of the IQE procedure on single-laboratory data sets that were designed to characterize an analytical method in the region of detection and quantitation. Despite the lack of comprehensive testing performed to date, however, EPA believes that the IQE procedure can be tested if sufficient resources are invested. In other words, the IQE meets the condition that it “can be” tested, but only partially meets the condition that it “has been” tested.

Condition 2: It has been subjected to peer review and publication. Although the IQE has not been published in the peer-reviewed scientific literature, the IQE has undergone extensive review and ballot by members of ASTM Committee D 19, many of whom are qualified peer reviewers. Therefore, although the IQE does not meet this condition in the sense of formal peer review and publication, EPA believes it does meet the intent of this condition (i.e., submission to scrutiny of the scientific community).

Condition 3: The error rate associated with the procedure is either known or can be estimated. In theory, an expert statistician could estimate the error rate of the IQE. However, the IQE procedure is extremely complex from an analytical chemistry and statistical perspective. As a result, it is unlikely that the error rate could be estimated by the staff of an environmental testing laboratory. Moreover, in attempting to follow the IQE procedure during this reassessment, EPA found the procedure to be highly subjective, particularly with respect to selection of an appropriate model. The subjective nature of the procedure is likely to yield different IQEs from the same data set, depending on the staff involved in analyzing the data and performing the calculations. (The likelihood of this problem is illustrated in Appendix C to this Assessment Document.) EPA believes such conditions make it difficult, if not impossible, to estimate the actual error associated with the IQE. Therefore, the IQE fails this condition.

Condition 4: Standards exist and can be maintained to control its operation. The IQE approach and procedure is supported by a published procedure (standard) to control its operation. The procedure gives the steps to be followed in determining the IQE and instructs the study supervisor how to gather the data and compute an IQE.

However, there are several "gray areas" in the published procedure. The most significant gray area is in model selection. The procedure provides insufficient guidance on the use of residual plots as a basis for selecting models and as a result, selection of the model may be very subjective, especially if the number of concentrations is low. The discussion of what model to use after rejecting the hybrid and linear models also is very vague. The exponential model is mentioned, as well as models with more than one coefficient. Much of the data evaluated by EPA tended to suggest the exponential model, based on the statistical tests discussed. However, those data have almost always shown residual “patterns” when using this model, which would then lead to consideration of other models. In addition, fitting the “constant model” is never discussed in detail. Most likely, this is done by simply calculating a mean (weighted if necessary) of the variances from the different concentrations, however such a calculation never explicitly stated.

As discussed under Condition 4 of Section 5.1.2.2.1 (scientific validity of the IDE procedure), EPA also is concerned about inconsistencies between the IDE and IQE that suggest conceptual problems with these standards. Finally, EPA observed that the IQE contains statistical errors that, if followed as written, could produce inaccurate IQE values. For example, the computations for weighted least squares

given in Table 1 of the procedure are incorrect. The formulae for the weighted means of the spike values and results given in Table 1 of D6512 would only be appropriate if the weighting were done based on the number of replicates per spike level, rather than on the estimated variance calculated using the chosen standard deviation model.

Based on these findings (along with those discussed under Criterion 2 below), EPA believes that, although the IQE is supported by a published procedure, the procedure is not sufficient to control operation of the IQE because of the high degree of subjectivity involved in implementing the procedure, statistical errors in the procedure, and internal inconsistencies with the IDE. Therefore, the IQE fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. The IQE fails this condition because it is familiar to, and has been accepted only by, a very narrow segment of the scientific community. Although the IQE has been approved by ASTM for more than 2 years, EPA has not found an IQE in the open literature or in an analytical method, including an ASTM method.

5.2.2.2.2 Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability.

The IQE procedure is designed to reflect expectations of interlaboratory performance, including routine variability. The procedure contains extensive instructions for dealing with unusual conditions, including sources of variability and outliers. Based on studies of the single-laboratory variant of the procedure in which the model selection proved to be highly subjective, EPA is skeptical about the procedure being able to demonstrate realistic expectations of laboratory and method performance.

The IQE procedure suggests attempting to fit study results to a constant, linear, or hybrid model. If all of these fail, the procedure suggests trying a different model, such as the exponential model. (The exponential model figures more prominently in the IDE procedure, where it is one of the three main models discussed, replacing the Rocke and Lorenzato model.) Although the exponential model may be appropriate for the IDE (which is not tied to a fixed RSD), it yields unacceptable results when applied to the IQE procedure. Under the exponential model, relative variability (standard deviation divided by the true concentration) is a parabolic function (i.e., as concentration increases, relative variability decreases down to a specific percentage, and then begins to increase). This is not realistic of laboratory and method performance. In addition, the exponential model will often result in having two possible values each for $IQE_{10\%}$, $IQE_{20\%}$, and $IQE_{30\%}$.

Another concern with the IQE procedure is that use of the non-mandatory appendices in ASTM D 6512 to determine the fit of a model may produce results that differ from those that would be obtained using the default procedures for testing model fit that are built into off-the-shelf statistical software.

Given the subjectivity and confusion involved in selecting the model, EPA tried using the same data set to calculate a single-laboratory variant of the IQE with each of the available models and found that the calculated IQEs varied widely when different models were used.

Based on the problems described above, EPA believes the IQE fails this criterion.

5.2.2.2.3 Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.

The IQE procedure is neither practical nor affordable in a single-laboratory context.. It is designed for use by an ASTM study supervisor or task manager and not as a procedure that a single

laboratory can use to evaluate method performance. EPA is aware that ASTM Committee D 19 is contemplating development of a within-laboratory quantitation estimate (WQE), but the WQE has not been approved through an ASTM ballot and therefore, it cannot be adequately evaluated at this time. The WQE may meet this criterion, but the IQE does not.

Regarding affordability, EPA estimates that the cost of implementing IQE procedure would be more than twice the cost of EPA's present implementation of the ML. The increased cost stems from the additional low-level data required to assure that variability versus concentration is being characterized in the region of detection and quantitation, challenges involved in applying the statistical procedures in the IQE, and because of the anticipated reanalysis and rework required if either the procedure failed to produce an IQE or if the resulting IQE failed to meet the specifications in the IQE procedure.

5.2.2.2.4 *Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.*

If the IQE were developed in an interlaboratory study that met the requirements of D 6512, the calculated IQE would likely be achievable by experienced staff in a well-operated laboratory. Therefore, the IQE passes this criterion. However, EPA also notes that although it passes the criterion, based on this assessment, EPA believes that it is very likely that the IQE may not identify the *lowest* concentration at which the signal is recognizable when the method is performed by experienced staff in a well-operated laboratory.

5.2.2.2.5 *Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government*

Although the IQE could be applied to some decisions to be made under CWA, it may not support decisions when pollutant levels need to be protective of human health and the environment because the IQE may be considerably higher than these levels. At best, the IQE only partially passes this criterion.

5.2.3 Assessment of the ACS Limit of Quantitation

The Limit of Quantitation (LOQ) was developed by the Committee on Environmental Improvement of the American Chemical Society (ACS) and published in the same two papers as the LOD.

5.2.3.1 Description of the ACS LOQ Approach and Procedure

The 1983 "*Principles*" define the LOQ as:

"... the level above which quantitative results may be obtained with a specified degree of confidence."

The same relationship used to define the LOD is used for the LOQ:

$$S_t - S_b \geq K_d \sigma$$

but the recommended minimal value for K_d be set at 10. Thus, the LOQ is 10σ above the gross blank signal, S_b . According to the 1983 publication, the LOQ corresponds to an uncertainty of $\pm 30\%$ ($10\sigma \pm$

3 σ). This uncertainty statement is based on σ equal to 10% of the LOQ. Other statements of uncertainty are, of course, possible using knowledge of σ and/or the RSD.

Neither the 1980 nor 1983 ACS publications provide a specific procedure for estimating the LOQ, nor do they provide a minimum number of observations needed to estimate the gross blank signal or the variability term σ_b .

5.2.3.2 Assessment of the ACS LOQ Against the Evaluation Criteria

The following five subsections discuss the ACS LOQ approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.3.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. Testing of the LOQ is hampered by 1) the lack of a supporting procedure for establishing an LOQ, and 2) its conceptual dependence on the variability of blank measurements. If the blank measurements fail to produce a response, it is impossible to calculate an LOQ because the standard deviation of zero is zero. One solution for testing the approach is to assume that the LOQ is functionally equivalent to the ML as the blank signal approaches zero. EPA believes this is a reasonable assumption, and therefore, that the ML procedure is a viable means for testing the LOQ approach. Therefore, the LOQ meets this condition.

Condition 2: It has been subjected to peer review and publication. The ACS LOQ definition was published in the peer-reviewed journal *Analytical Chemistry* in 1980 and 1983. Therefore, the ACS LOQ meets this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. The definition of the LOQ specifically estimates the uncertainty associated with a concentration at the LOQ as $\pm 30\%$ based on 10% RSD. Other valid statements in terms of %RSD may be made based on study requirements, policy judgments and/or specific results. For example, the estimate of an uncertainty of $\pm 30\%$ based on 10% RSD is inconsistent with EPA and ISO/IUPAC estimations that place the uncertainty at $\pm 20\%$ (at $\pm 2\sigma$), and is inconsistent with the Episode 6000 data that place the median RSD at 7% and therefore, the $\pm 2\sigma$ uncertainty at approximately $\pm 14\%$.

Condition 4: Standards exist and can be maintained to control its operation. The ACS LOQ lacks a clearly defined procedure for estimating the important terms required to derive it. Although it may be possible to derive ACS LOQ values from data used to derive EPA MDL values, there is no discussion of using replicate blanks, replicate spiked samples, or a minimum recommendation for the number of replicates. Therefore, the ACS LOQ fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Because the ACS does not develop and publish reference analytical methods, it is difficult to determine the degree of acceptance of the LOQ. EPA has not investigated the numbers of papers published in ACS journals that include LOQ values, but EPA's literature search for detection and quantitation approaches did not uncover a large number of citations that promote the LOQ in particular.

5.2.3.2.2 *Criterion 2: The approach should address demonstrated expectations of laboratory and method performance, including routine variability*

The LOQ approach is designed to address demonstrated expectations of laboratory and method performance, including routine variability, and therefore, it appears to meet this criterion. Unfortunately, ACS has not published a procedure to implement the approach. In other words, the LOQ addresses demonstrated expectations of laboratory and method performance in theory, but in practice, provides no direct means for performing these demonstrations. Therefore, EPA believes the ACS LOQ only partially meets this criterion.

5.2.3.2.3 *Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The ACS LOQ approach is not supported by a clearly defined procedure for establishing the LOQ. Therefore, it fails this criterion.

5.2.3.2.4 *Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.*

Given the relationship of the ACS LOQ to the ML, EPA believes the LOQ meets this criterion for the reasons outlined in Section 5.2.1.2.4, which discusses EPA's assessment of the ML against Criterion 4 for evaluating quantitation limit approaches.

5.2.3.2.5 *Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.*

In the absence of a procedure for determining LOQ values, the ACS LOQ fails to meet this criterion because it cannot be used in a regulatory context. The LOQ passes this criterion only if it is assumed to be functionally equivalent to the ML (i.e., the ML procedure is used to establish an LOQ).

5.2.4 Assessment of the IUPAC/ISO Limit of Quantitation

A similar LOQ approach was developed by IUPAC/ISO and published in the same papers as the CRV and MDV (see Sections 5.1.4 and 5.1.5).

5.2.4.1 Description of the ISO/IUPAC LOQ Approach

The 1995 "Recommendations" define the LOQ as:

"... the ability of a CMP [chemical measurement process] to adequately 'quantify' an analyte. The ability to quantify is generally expressed in terms of the signal or analyte (true) value that will produce estimates having a specified relative standard deviation (RSD), commonly 10 %."

The relationship used to define the LOQ is:

$$L_Q = K_Q \times \sigma_Q$$

The recommended value for K_Q is 10. Thus, the LOQ is 10σ above the blank signal, σ_Q .

5.2.4.2 Assessment of the IUPAC/ISO LOQ Against the Evaluation Criteria

The following five subsections discuss the IUPAC/ISO LOQ approach and procedure in the context of the five evaluation criteria that concern detection limit approaches (i.e., Criteria 1-3, and Criteria 5 and 6).

5.2.4.2.1 Criterion 1: *The detection and quantitation limit approaches should be scientifically valid.*

Condition 1: It can be (and has been) tested. Testing of the IUPAC/ISO LOQ is hampered by 1) the lack of a supporting procedure for establishing and LOQ, and 2) it's conceptual dependence on the variability of blank measurements. If the blank measurements fail to produce a response, it is impossible to calculate an LOQ because the standard deviation of zero is zero. One solution for testing the approach is to assume that the ISO/IUPAC LOQ is functionally equivalent to the ML as the blank signal approaches zero. EPA believes this is a reasonable assumption, and that the ML procedure is a viable means for testing the LOQ approach. Therefore, the ISO/IUPAC LOQ meets this condition.

Condition 2: It has been subjected to peer review and publication. The IUPAC/ISO LOQ definition has been published by Currie in the peer-reviewed journals *Pure and Appl. Chem.* in 1995; in *Anal. Chim. Acta* in 1999, in *Chemometrics and Intelligent Lab Systems* in 1997; and in *J. Radioanal. and Nuclear Chem.* in 2000. Therefore, the IUPAC/ISO LOQ meets this condition.

Condition 3: The error rate associated with the procedure is either known or can be estimated. EPA used data generated in the Episode 6000 study to estimate the error rate associated with the LOQ. The Episode 6000 results show that the median error across all analytes and analytical techniques at 10σ is approximately $\pm 14\%$ with approximately 95% confidence.

Condition 4: Standards exist and can be maintained to control its operation. The IUPAC/ISO LOQ lacks a clearly defined procedure for estimating the important terms required to derive it. Although it may be possible to derive IUPAC/ISO LOQ values from data used to derive EPA MDL values, there is no discussion of using replicate blanks, replicate spiked samples, or a minimum recommendation for the number of replicates. Therefore, EPA believes that the IUPAC/ISO LOQ fails this condition.

Condition 5: It has attracted widespread acceptance within a relevant scientific community. Acceptance by the scientific community is not known. Acceptance would be indicated by use of the LOD in ISO methods. EPA did not perform a search of ISO methods because of copyright restrictions. However, EPA's literature search for detection and quantitation approaches in the open technical literature did not uncover a large number of citations that reference the LOQ. Therefore, it is difficult to determine if the ISO/IUPAC LOQ meets this condition.

5.2.4.2.2 Criterion 2: *The approach should address demonstrated expectations of laboratory and method performance, including routine variability.*

The most recent publication on the IUPAC/ISO LOQ (*J. Radioanal. and Nuclear Chem.*, op. cit.) provides insight into this issue through measurements of ^{14}C by accelerator mass spectrometry. Therefore, EPA believes that the IUPAC/ISO LOQ passes this criterion for at least some measurement techniques.

5.2.4.2.3 *Criterion 3: The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.*

The ISO/IUPAC LOQ approach is not supported by a clearly defined procedure for establishing the LOQ. Therefore, it fails this criterion.

5.2.4.2.4 *Criterion 5: The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.*

Given the relationship of the IUPAC/ISO LOQ to the ML, EPA believes that the LOQ satisfies this criterion for the reasons outlined in Section 5.2.1.2.4, which discusses EPA's assessment of the ML against Criterion 4 for evaluating quantitation limit approaches.

5.2.4.2.5 *Criterion 6: Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government*

In the absence of a procedure for determining LOQ values, the ISO/IUPAC LOQ fails to meet this criterion because it cannot be used in a regulatory context. The ISO/IUPAC LOQ passes only if the ML procedure is used to establish an LOQ.

Table 5-1. Assessment of Detection Limit Approaches Against Evaluation Criteria					
Evaluation Criteria	MDL	IDE	ACS LOD	ISO/IUPAC CRV	ISO/IUPAC MDV
<p>The detection limit approach should be scientifically valid:</p> <ul style="list-style-type: none"> • It can be (and has been tested) • Has undergone peer review and publication • Has an error rate that is known or can be estimated • Has standards that can be maintained to control its operation • Has achieved widespread acceptance in a relevant scientific community 	Meets all 5 conditions for scientific validity with slight modifications noted to clarify understanding of error rate.	<p>Meets 1, partially meets 1, and fails 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • Can be, but has not been fully tested (partial) • Subjectivity makes calculation of error rate impossible (fails) • Has a standard but, due to the high degree of subjectivity, errors, and conceptual inconsistency, it is unlikely to control its operation (fails) • Is familiar to and accepted by a very narrow segment of the scientific community (fails) 	<p>Meets 4 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • No standards exist to control its operation 	<p>Meets 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • No standards exist to control its operation • Degree of acceptance is unclear 	<p>Meets 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • No standard exist to control its operation • Degree of acceptance is unclear
The approach should address demonstrated expectations of laboratory and method performance, including routine variability.	Can meet this criterion if properly applied.	Conceptually passes this criterion, but fails in practice due to problems with model selection	Partially meets the criterion. Approach meets the criterion but no procedure for implementing the approach is given. Passes the criterion only if equivalency to the MDL is assumed.	Partially meets this criterion. Approach meets the criterion but no procedure for implementing the approach is given. Passes the criterion only if equivalency to the MDL is assumed.	Could be used in planning method development and evaluation studies as recommended but not in operational detection decision making.
The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.	Meets this criterion. Procedure can be performed by a single laboratory during a single shift, or for method development by multiple labs in a single shift	Fails this criterion. Requires interlaboratory study involving a reference lab or coordinating body, a minimum of 6 complete data sets, and a skilled statistician. The cost of implementing this procedure would exceed most method development budgets.	Fails this criterion. No procedure provided.	Fails this criterion. No procedure provided.	Fails this criterion. No procedure provided.

Table 5-1. Assessment of Detection Limit Approaches Against Evaluation Criteria					
Evaluation Criteria	MDL	IDE	ACS LOD	ISO/IUPAC CRV	ISO/IUPAC MDV
The detection level approach should identify the signal or estimated concentration at which there is 99% confidence that the substance is actually present when the analytical method is performed by experienced staff in a well-operated laboratory.	Meets this criterion.	When the allowance for false negatives and for prediction and tolerance are taken into account, the resulting detection limit (IDE) is raised to the point at which detection probability is estimated to be greater than 99.999999%; this yields numerical values that have no practical meaning as a detection standard. Therefore, the IDE fails this criterion.	Meets this criterion.	Meets this criterion.	The MDV is a true concentration value not used in the actual detection decision. Does not meet the criterion.
Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.	Meets this criterion.	At best, only partially passes this criterion. Not likely to meet this criterion in instances in which a compliance limit is close to a detection limit determined by a procedure such as the MDL.	In the absence of a procedure for determining LOD values, fails to meet this criterion.	In the absence of a procedure for determining CRV values, fails to meet this criterion.	In the absence of a procedure for determining MDV values, fails to meet this criterion.

Table 5-2. Assessment of Quantitation Limit Approaches Against Evaluation Criteria				
Evaluation Criteria	ML	IQE	ACS LOQ	ISO/IUPAC LOQ
<p>The quantitation limit approach should be scientifically valid.</p> <ul style="list-style-type: none"> • It can be (and has been tested) • Has undergone peer review and publication • Has an error rate that is known or can be estimated • Has standards that can be maintained to control its operation • Has achieved widespread acceptance in a relevant scientific community 	<p>Meets all 5 conditions for scientific validity, though slight modification to the definition is suggested to improve operation when other than 7 replicates are used to estimate the ML.</p>	<p>Meets 1 condition, partially meets 1 condition, and fails 3 conditions.</p> <ul style="list-style-type: none"> • Can be, but has not been fully tested (partial) • Error rate cannot be estimated due to problems with the procedure (fail) • Standards are not likely to control its operation (fail) • Has not achieved widespread acceptance (fail) 	<p>Meets 3 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • Lacks a standard to control its operation • Difficult to determine the degree of acceptance 	<p>Meets 4 of the 5 conditions for scientific validity.</p> <ul style="list-style-type: none"> • Lacks a standard to control its operation • Difficult to determine the degree of acceptance
<p>The approach should address demonstrated expectations of laboratory and method performance, including routine variability.</p>	<p>Meets this criterion. Procedure can be performed by a single laboratory during a single shift, or for method development by multiple labs in a single shift.</p>	<p>Fails this criterion due to subjectivity, errors, and theoretical inconsistencies in the procedure.</p>	<p>Partially meets this criterion. The approach is designed to address these expectations but in practice, there is no procedure for performing such demonstrations.</p>	<p>Meets this criterion.</p>
<p>The approach should be supported by a practical and affordable procedure that a single laboratory can use to evaluate method performance.</p>	<p>Meets this criterion.</p>	<p>Fails this criterion. Requires interlaboratory study involving a reference lab or coordinating body, 6 complete data sets, and a highly skilled statistician. The cost of implementing this procedure would exceed most method development budgets.</p>	<p>Fails this criterion.</p>	<p>Fails this criterion.</p>
<p>The quantitation limit approach should identify the concentration that gives a recognizable signal that is consistent with the capabilities of the method when a method is performed by experienced staff in well-operated laboratories.</p>	<p>Meets this criterion.</p>	<p>Meets this criterion, but is not likely to estimate the <i>lowest</i> level at which reliable measurements can be made by an experienced analyst in a well operated lab</p>	<p>Meets this criterion.</p>	<p>Meets this criterion.</p>

Table 5-2. Assessment of Quantitation Limit Approaches Against Evaluation Criteria				
Evaluation Criteria	ML	IQE	ACS LOQ	ISO/IUPAC LOQ
Detection and quantitation approaches should be applicable to the variety of decisions made under the Clean Water Act, and should support state and local obligations to implement measurement requirements that are at least as stringent as those set by the Federal government.	Meets this criterion.	At best, only partially passes this criterion. Fails for those instances in which the IQE limit is greater than an effluent limit or water quality-based limit.	Fails this criterion. In the absence of a procedure for determining ACS LOQ values, the ACS LOQ cannot be used in a regulatory context.	Fails this criterion. In the absence of a procedure for determining LOQ values, the ISO/IUPAC LOQ cannot be used in a regulatory context.

This chapter summarizes the results of EPA's assessment of detection and quantitation limit approaches. This assessment, which is detailed in the previous five chapters, was based on:

- Identification of relevant approaches to include in the assessment (Chapter 2),
- Identification of issues that may be relevant to the assessment from an analytical chemistry, statistical, or regulatory perspective (Chapter 3),
- Development of criteria that reflect EPA's views concerning these issues (Chapter 4) and form the primary basis for evaluating the ability of each approach to meet EPA needs under the Clean Water Act,
- Assessment of how well each approach met the evaluation criteria (Chapter 5), and,
- Use of real-world data to evaluate both the theoretical and practical limitations of each approach (Appendices B and C).

EPA evaluated four sets of detection and quantitation limit approaches advanced by EPA, ASTM International, ACS, and both ISO and IUPAC. Each approach was assessed against the suite of criteria described in Chapter 4. The EPA approaches (i.e., the MDL and ML) and the ASTM International approaches (i.e., the IDE and IQE) were supported by clearly defined procedures for implementing the approach. Neither the ACS nor the ISO/IUPAC approaches are supported by detailed procedures for implementation; this lack of supporting procedures was reflected in the outcome of EPA's overall assessment.

After evaluating each approach against each of the evaluation criteria, EPA found that 1) no single pair of detection and quantitation limit approaches perfectly meets EPA's criteria, 2) the MDL and ML are closest to meeting EPA's criteria, and 3) minor revisions and clarifications to the MDL and ML would allow both approaches to fully meet the Agency's needs under the CWA.

EPA also found that, although the IDE and IQE procedures may be acceptable for planning and implementing interlaboratory studies to develop and validate analytical methods, there are a number of difficulties with these procedures that make them unsuitable as the primary means of establishing sensitivity under the Clean Water Act. In particular, the IDE is analogous by definition and formulaic construction to the "*Detection Limit*" defined by Currie (1968, 1995), while it is Currie's "*Critical Value*" approach that is most relevant to Agency needs under the CWA. Currie (1995) states that the decision "detected" or "not detected" is made by comparison of the estimated quantity or measured value with the *critical value*. Currie describes his "*Detection Limit*" as a true concentration that has a high probability of generating measured values that exceed the *critical value*, and states that the single most important application of the detection limit is for planning and evaluation of measurement procedures and that the *detection limit*:

"...allows one to judge whether the CMP (Chemical Measurement Process) under consideration is adequate for detection requirements. This is in sharp contrast to application of the critical value for decision making, given the result of a measurement."

It is important to note that the formulation of the MDL is analogous to the Currie critical value, and as such, is intended to be used to make detection decisions in the manner described by Currie (i.e., the MDL is designed and used to make the decision of "detected" or "not detected"). EPA believes that form of detection decision best supports the use of "detection limits" under CWA programs.

In contrast, although the IDE is intended to be used in a manner analogous to Currie's *critical value* (i.e., to make the decision of "detected" or "not detected"), it is, by definition and design, functionally analogous to Currie's *detection limit* (i.e., it identifies a concentration that will have a high probability of generating measured values that exceed the critical value). (See Chapter 2, Section 2.1 for a discussion of Currie's *critical value* and *detection limit*).

Other drawbacks with the ASTM International approach include the complexities of the IDE and IQE procedures, along with their inability to address individual laboratory performance. Despite these limitations, however, EPA believes the IDE and IQE can be used to establish sensitivity for certain applications. For example, consider the theoretical situation of an ASTM method for the determination of an analyte regulated under the NPDES program that uses the IDE or IQE to describe method sensitivity *and* for which the value of the IDE or IQE was below the relevant criterion or regulatory limit. EPA would evaluate the overall performance of such a method for approval at 40 CFR part 136, despite the fact that the method did not contain an MDL determined using the procedure described in 40 CFR part 136, Appendix B. (See Chapter 3, Section 3.2.8 for a more in-depth discussion of using alternative procedures to establish sensitivity.)

EPA's assessment of the theoretical and practical applications of each detection and quantitation approach (see Appendices B and C) is summarized in Exhibit 6-1. This exhibit suggests that no approach produces the "right" answer, and that different approaches produce different detection and quantitation limits. Observed differences are largely due to different sources of variability accounted for among the approaches.

As part of this assessment, EPA identified the need for approaches that can support CWA programs, including:

- method performance verification at a laboratory,
- method development and promulgation,
- National Pollutant Discharge Elimination System (NPDES) applications,
- non-regulatory studies and monitoring,
- descriptive versus prescriptive uses of lower limits to measurement, and
- use of a pair of related detection and quantitation procedures in all OW applications

EPA has concluded that the MDL and ML can meet all of these applications and that the addition of a scope and application section to the procedure would help clarify use of the MDL for these applications. However, as noted in Chapter 3, outside organizations use different detection and quantitation approaches that meet their own needs. Given EPA's diverse needs and desire to encourage the development of improved measurement techniques, EPA does not believe it is necessary or appropriate to require the exclusive use of the MDL and ML approaches in CWA programs. As indicated above, EPA would allow use of alternative detection and quantitation procedures to establish detection and quantitation limits in an analytical method, provided that the resulting detection and quantitation limits meet the sensitivity needs for the specific application.

Exhibit 6-1: Theoretical and Practical Application of Each Approach

Finding 1: Each approach yields different values.

Detection Limit Approaches

- The EPA MDL and ACS LOD approaches, which are functionally analogous, produced detection limits that are a median of 1.25 times higher than the limits produced by the CRV advanced by ISO and IUPAC (Appendix C of this document).
- The Minimum Detectable Value (MDV) advanced by ISO and IUPAC produced detection limits that are a median of 1.2 times higher than the limits produced by the MDL and LOD approaches (Appendix C of this document).
- A single-laboratory variant of the IDE (the IDE has been advanced by ASTM International) produced detection limits that are a median of 2.9 times higher than the median limits produced by the MDL and LOD approaches (Appendix C of this document). This result is not surprising given that the IDE is functionally analogous to Currie's *detection level*, while the MDL and LOD are analogous to Currie's *critical value*.

Quantitation Limit Approaches

- The EPA ML and the functionally equivalent ACS LOQ produced quantitation limits that are a median of 1.1 times higher than the limits produced by the LOQ approach advanced by ISO and IUPAC (Appendix C of this document).
- A single-laboratory variant of the IQE (the IQE has been advanced by ASTM International) produced median quantitation limits that are equivalent to the median limits produced by the EPA ML and ACS LOQ approaches (Appendix C of this document).

Finding 2: More than the 5 levels specified by ASTM are required to produce a reliable IDE and IQE

- EPA found that the IDEs produced with a subset of data generated from the minimum of 5 concentrations recommended in the IDE procedure differed widely from the IDEs produced with a larger set of data involving 16 concentrations (which included the subset of 5 concentrations) (Appendix C of this document).
- Findings suggest that more than 5 concentrations are needed to produce a reliable IDE, due to the limited power of the statistical tests for significant model parameters and the difficulty of drawing conclusions based on residual plots with only 5 points (Appendix C of this document).
- Parallel reasoning can be applied to the IQE based on its similarity to the IDE.

Finding 3: The ML procedure yields quantitation limits that are generally in the range of the 10% RSD intended in the ML (and the functionally analogous ACS LOQ) approach.

- EPA calculated the uncertainty associated with replicate measurements made at the ML for a large number of analytes and techniques (Appendix C of this document).
- EPA found that on average, across all techniques tested, the RSD of replicate measurements at the ML was approximately 7%. Median RSDs calculated for each multi-analyte method ranged from 6 - 14%, and RSD values calculated for each single-analyte method ranged from 4 - 16% (Appendix C of this document).

Finding 4: No single model adequately predicts the behavior of all analytes and all methods across the measurement range.

- EPA produced graphs representing hundreds of analyte/method combinations. Selection of an appropriate model based on these graphs is highly subjective, at best, due to the lack of clear patterns and the residuals observed with each model applied (Chapter 3, Section 3.3, and Appendix B of this document).
- The IDE and IQE are the only approaches other than the MDL and ML that are supported by a procedure for their implementation. The IDE and IQE procedures rely heavily on model selection, and the degree of subjectivity involved in selecting these models makes implementation of the IDE and IQE difficult (Chapter 5, Sections, 5.1.2 and 5.2.2, and the third conclusion in Appendix C).

Finding 5: Use of a recovery correction when establishing detection and quantitation limits may not be appropriate.

- EPA found that using a regression to estimate a recovery correction at zero concentration causes great swings in the resulting detection and quantitation limits (Appendix C of this document).
- Use of a recovery-correction procedure also can result in 'double-correcting' for recovery because 1) nearly all methods already contain specifications for acceptable recovery performance, and 2) some methods include recovery correction in the computation of sample results (Chapter 3, Section 3.1.4).

Exhibit 6-2: Summary of Recommended Modifications to the MDL and ML procedures

EPA believes that the following revisions and clarifications to the MDL and ML would allow these procedures to fully meet the Agency's needs under the CWA.

- Refine the definition of the MDL to make it more consistent with the MDL procedure and note the functional analogy of the MDL with the "critical value" described by Currie (1968 and 1995) and with the "limit of detection" (LOD) described by the American Chemical Society in 1980 and 1983 (Chapter 5, Section 5.2.1.1.1).
- Expand the Scope and Application discussion to acknowledge that there are a variety of purposes and analytical methods for which the MDL procedure may be employed and to provide examples of common uses of the MDL procedure (i.e., demonstrating laboratory capability with a particular method; monitoring trends in laboratory performance; characterizing method sensitivity in a particular matrix; and establishing an MDL for a new or revised method for nationwide use).
- Clarify the considerations for estimating the detection limit in Step 1 of the current MDL procedure, and suggest that the method-specified MDL can be used as the initial estimate when performing an MDL study to verify laboratory performance or to demonstrate that the MDL can be achieved in a specific matrix (Chapter 5, Section 5.2.1.1.1).
- Revise the specifications for establishing the test concentration range (i.e., determining the spike levels) in Section 3.1 according to the intended application of the MDL as follows: 1) if verifying a published MDL, the test concentration should be no more than five times the published MDL; 2) if verifying an MDL to support a regulatory objective or the objective of a study or program, the test concentration should be no more than one third the compliance or target limit; 3) if determining an MDL for a new or revised method, the test concentration should be no more than five times the estimated detection limit; and 4) if performing an iteration, the test concentration should be no more than five times the MDL determined in the most recent iteration.
- Delete the calculation of a 95% confidence interval estimate for the MDL from Step 6. EPA has determined that these calculations are neither routinely performed by laboratories, nor are the results employed by regulatory agencies, including EPA.
- Revise Step 7 to 1) require that the iterative procedure be used to verify the reasonableness of the MDL when developing an MDL for a new or revised method or when developing a matrix-specific MDL, but that it remain optional when verifying a method-, matrix-, program-, or study-specific MDL, and 2) provide specific instructions on how to assess the reasonableness of an MDL used to verify laboratory performance (Chapter 5, Section 5.2.1.1.1).
- Add a new Step 8 to the MDL procedure to address the treatment of suspected outliers (Chapter 5, Section 5.2.1.1.1).
- Delete the discussion of analysis and use of blanks included in Section 4(a) of the current MDL procedure. The current discussion applies to methods in which a blank measurement is required to calculate the measured level of an analyte; it requires separate measurements of blank samples for each MDL sample aliquot analyzed and subtraction of the average result of the blank samples from each respective MDL sample measurement. Deletion of this discussion recognizes that subtraction of a single (or average) blank sample result from the result for each MDL sample would not change the standard deviation and thus, would have no effect on the resulting MDL. Although EPA believes laboratories would be prudent to analyze method blanks for assessing potential contamination, EPA also believes that requiring analysis of method blanks or subtraction of method blank results during MDL determinations is unnecessarily burdensome.
- Revise the optional pre-test described in Section 4(b) of the current MDL procedure to provide criteria that allow the analyst to determine if the test samples are the desirable range.
- Improve overall readability and understanding of the MDL procedure through editorial changes to the specific numbering scheme, the addition of clearer titles to some of the steps, and minor clarifications.
- Clarify the ML to emphasize its relationship to Currie's Quantitation Limit and ACS' Limit of Quantitation (LOQ)
- Clarify the ML procedure to address the use of other than seven replicates for determination of the MDL and ML.